

Understanding Complex Systems

Springer :
COMPLEXITY

Jan Treur

Network- Oriented Modeling

Addressing Complexity of Cognitive,
Affective and Social Interactions



Springer

Springer Complexity

Springer Complexity is an interdisciplinary program publishing the best research and academic-level teaching on both fundamental and applied aspects of complex systems—cutting across all traditional disciplines of the natural and life sciences, engineering, economics, medicine, neuroscience, social and computer science.

Complex Systems are systems that comprise many interacting parts with the ability to generate a new quality of macroscopic collective behavior the manifestations of which are the spontaneous formation of distinctive temporal, spatial or functional structures. Models of such systems can be successfully mapped onto quite diverse “real-life” situations like the climate, the coherent emission of light from lasers, chemical reaction-diffusion systems, biological cellular networks, the dynamics of stock markets and of the internet, earthquake statistics and prediction, freeway traffic, the human brain, or the formation of opinions in social systems, to name just some of the popular applications.

Although their scope and methodologies overlap somewhat, one can distinguish the following main concepts and tools: self-organization, nonlinear dynamics, synergetics, turbulence, dynamical systems, catastrophes, instabilities, stochastic processes, chaos, graphs and networks, cellular automata, adaptive systems, genetic algorithms and computational intelligence.

The three major book publication platforms of the Springer Complexity program are the monograph series “Understanding Complex Systems” focusing on the various applications of complexity, the “Springer Series in Synergetics”, which is devoted to the quantitative theoretical and methodological foundations, and the “Springer Briefs in Complexity” which are concise and topical working reports, case studies, surveys, essays and lecture notes of relevance to the field. In addition to the books in these two core series, the program also incorporates individual titles ranging from textbooks to major reference works.

Editorial and Programme Advisory Board

Henry Abarbanel, Institute for Nonlinear Science, University of California, San Diego, USA

Dan Braha, New England Complex Systems Institute and University of Massachusetts Dartmouth, USA

Péter Érdi, Center for Complex Systems Studies, Kalamazoo College, USA and Hungarian Academy of Sciences, Budapest, Hungary

Karl Friston, Institute of Cognitive Neuroscience, University College London, London, UK

Hermann Haken, Center of Synergetics, University of Stuttgart, Stuttgart, Germany

Viktor Jirsa, Centre National de la Recherche Scientifique (CNRS), Université de la Méditerranée, Marseille, France

Janusz Kacprzyk, System Research, Polish Academy of Sciences, Warsaw, Poland

Kunihiko Kaneko, Research Center for Complex Systems Biology, The University of Tokyo, Tokyo, Japan

Scott Kelso, Center for Complex Systems and Brain Sciences, Florida Atlantic University, Boca Raton, USA

Markus Kirkilionis, Mathematics Institute and Centre for Complex Systems, University of Warwick, Coventry, UK

Jürgen Kurths, Nonlinear Dynamics Group, University of Potsdam, Potsdam, Germany

Ronaldo Menezes, Department of Computer Science, Florida Institute of Technology, Melbourne, FL, USA

Andrzej Nowak, Department of Psychology, Warsaw University, Poland

Hassan Qudrat-Ullah, School of Administrative Studies, York University, Toronto, ON, Canada

Peter Schuster, Theoretical Chemistry and Structural Biology, University of Vienna, Vienna, Austria

Frank Schweitzer, System Design, ETH Zurich, Zurich, Switzerland

Didier Sornette, Entrepreneurial Risk, ETH Zurich, Zurich, Switzerland

Stefan Thurner, Section for Science of Complex Systems, Medical University of Vienna, Vienna, Austria

Understanding Complex Systems

Founding Editor: S. Kelso

Future scientific and technological developments in many fields will necessarily depend upon coming to grips with complex systems. Such systems are complex in both their composition – typically many different kinds of components interacting simultaneously and nonlinearly with each other and their environments on multiple levels – and in the rich diversity of behavior of which they are capable.

The Springer Series in Understanding Complex Systems series (UCS) promotes new strategies and paradigms for understanding and realizing applications of complex systems research in a wide variety of fields and endeavors. UCS is explicitly transdisciplinary. It has three main goals: First, to elaborate the concepts, methods and tools of complex systems at all levels of description and in all scientific fields, especially newly emerging areas within the life, social, behavioral, economic, neuro- and cognitive sciences (and derivatives thereof); second, to encourage novel applications of these ideas in various fields of engineering and computation such as robotics, nano-technology, and informatics; third, to provide a single forum within which commonalities and differences in the workings of complex systems may be discerned, hence leading to deeper insight and understanding.

UCS will publish monographs, lecture notes, and selected edited contributions aimed at communicating new findings to a large multidisciplinary audience.

More information about this series at <http://www.springer.com/series/5394>

Jan Treur

Network-Oriented Modeling

Addressing Complexity of Cognitive,
Affective and Social Interactions



Springer

Jan Treur
Department of Computer Science,
Behavioural Informatics Group
VU University Amsterdam
Amsterdam, Noord-Holland
The Netherlands

ISSN 1860-0832	ISSN 1860-0840 (electronic)
Understanding Complex Systems	
ISBN 978-3-319-45211-1	ISBN 978-3-319-45213-5 (eBook)
DOI 10.1007/978-3-319-45213-5	

Library of Congress Control Number: 2016948267

© Springer International Publishing Switzerland 2016

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

During a sabbatical period in 2015 I decided to start working on a book on the Network-Oriented Modeling approach developed over the past years and which has turned out useful in modeling complex-integrated individual and social human processes in the form of networks. The decision to spend a considerable amount of time on such an enterprise led to further reflection on the modeling approach, and its presentation and positioning.

This book has been written with a multidisciplinary audience in mind without assuming much prior knowledge. In principle, the detailed explanations in the book allow it to be used as an introduction in Network-Oriented Modeling for multidisciplinary Master and Ph.D. students. In particular, this implies that, although some more technical mathematical and formal logical aspects have also been addressed, they have been kept to the minimum and are presented in a concise manner in Part IV. They can be skipped if not needed. Much of the material in this book has been and is being used in teaching multidisciplinary courses for undergraduate and graduate students, and based on these experiences, the presentation has been adapted to suit requirements even better. Sometimes there is some overlap between chapters, and this has been done on purpose in order to make it easier to read each chapter separately. Lecturers can contact me to receive additional material such as slides, assignments and software.

The content of the book has benefited much from cooperation with students and (past and current) members of the Behavioural Informatics Group (formerly the Agent Systems Group) at the VU University in Amsterdam (Vrije Universiteit Amsterdam). In the discussion section in each of the chapters, specific publications and authors related to the material presented in the chapter are mentioned.

Amsterdam, The Netherlands
June 2016

Jan Treur

Contents

Part I Network-Oriented Modeling: Introduction

1	Network-Oriented Modeling and Its Conceptual Foundations	3
	An Introduction	3
1.1	Introduction	3
1.2	Addressing Human Complexity by Separation Assumptions	4
1.3	Addressing Complexity by Interaction in Networks Instead of by Separation	11
1.4	Network-Oriented Modeling	14
1.5	The Dynamic Computational Modeling Perspective	16
1.6	Network-Oriented Modeling Based on Temporal-Causal Networks	18
1.7	Scope of Applicability and Achievements.	22
1.8	Overview of the Book	23
	References.	29
2	A Temporal-Causal Network Modeling Approach	35
	With Biological, Neurological and Social Processes as Inspiration . . .	35
2.1	Introduction	35
2.2	Modeling Complex Processes by Temporal-Causal Networks	40
2.3	Exploiting Knowledge About Physical and Biological Mechanisms in Modeling	43
2.3.1	Addressing Complexity by Higher Level Models Based on Knowledge from Computer Science.	43
2.3.2	Addressing Complexity by Higher Level Models Based on Knowledge from Neuroscience	44
2.4	Conceptual Representation of a Temporal-Causal Network Model	45

2.4.1	Conceptual Representations of a Temporal-Causal Network Model	47
2.4.2	More Specific Examples of Conceptual Representations of Temporal-Causal Network Models.	49
2.5	Numerical Representation of a Temporal-Causal Network Model	58
2.5.1	The Systematic Transformation from Conceptual to Numerical Representation	59
2.5.2	Illustration of the Transformation for the Example of Fig. 2.10.	64
2.5.3	Illustration of the Modeling Perspective for a Social Contagion Process	66
2.6	Standard Combination Functions	69
2.6.1	Basic Standard Combination Functions	69
2.6.2	Building More Complex Standard Combination Functions.	72
2.7	Properties for Combination Functions.	77
2.8	Applying Computational Methods to Model Representations.	81
2.9	Applicability of the Modeling Perspective	85
2.9.1	The State-Determined System Assumption	85
2.9.2	State-Determined Systems and First-Order Differential Equations	86
2.9.3	State-Determined Systems and Modeling Based on Temporal-Causal Networks	88
2.10	Modeling Adaptive Processes by Adaptive Temporal-Causal Networks	92
2.11	Discussion.	99
	References.	100

Part II Emotions All the Way

3	How Emotions Come in Between Everything	105
	Emotions Serving as Glue in All Mental and Social Processes.	105
3.1	Introduction	105
3.2	Generating Emotional Responses and Feelings	107
3.3	Emotion Regulation	111
3.4	Interaction Between Cognitive and Affective States	114
3.5	Emotion-Related Valuing in Decision-Making	118
3.6	Emotions and Social Contagion	119
3.7	Discussion.	120
	References.	121

4	How Do You Feel Dreaming.	125
	Using Internal Simulation to Generate Emotional Dream Episodes. . . .	125
4.1	Introduction	125
4.2	Memory Elements, Emotions and Internal Simulation in Dreaming	126
4.3	A Temporal-Causal Network Model Generating Dream Episodes	128
4.4	Simulations of Example Dream Scenarios	133
4.5	Relations to Neurological Theories and Findings	136
4.6	Discussion.	137
	References.	138
5	Dreaming Your Fear Away	141
	Fear Extinction Learning During Dreaming	141
5.1	Introduction	141
5.2	An Adaptive Temporal-Causal Network Model for Fear Extinction Learning.	142
5.2.1	Conceptual Representation of the Adaptive Network Model	142
5.2.2	Numerical Representation of the Adaptive Network Model	146
5.3	Simulations of Fear Extinction Learning in Dream Scenarios	148
5.4	Relating the Adaptive Temporal-Causal Network Model to Neurological Theories.	152
5.5	Discussion.	153
	References.	154
6	Emotions as a Vehicle for Rationality in Decision Making	157
	Experiencing Emotions for Decisions Based on Experience	157
6.1	Introduction	157
6.2	The Adaptive Temporal-Causal Network Model for Decision Making.	159
6.3	Simulation Results for a Deterministic World.	168
6.4	Simulation Results for a Stochastic World	171
6.5	Simulation Results for a Changing Stochastic World	172
6.6	Evaluating the Adaptive Temporal-Causal Network Model on Rationality	175
6.7	Discussion.	178
	References.	179

Part III Yourself and the Others

7	From Mirroring to the Emergence of Shared Understanding and Collective Power	183
	Biological and Computational Perspectives on the Emergence of Social Phenomena	183
7.1	Introduction	183
7.2	Mirror Neuron Activation and Internal Simulation	185
7.2.1	The Discovery of Mirror Neurons	185
7.2.2	Neurons for Control and Self-other Distinction	186
7.2.3	Generating Emotions and Feelings by Internal Simulation: As-if Body Loops	187
7.2.4	Mirroring Process: Mirror Neuron Activation and Internal Simulation	187
7.2.5	Development of the Discipline Social Neuroscience	192
7.3	The Emergence of Shared Understanding	193
7.3.1	The Emergence of Shared Understanding for External World States	194
7.3.2	The Emergence of Shared Understanding for Internal Mental States	195
7.4	The Emergence of Collective Power	197
7.4.1	The Emergence of Collective Action Based on Mirroring	197
7.4.2	The Role of Feelings and Valuing in the Emergence of Collective Action	199
7.5	Integration of External Effects and Internal Processes	200
7.6	Abstraction of Complex Internal Temporal-Causal Network Models	202
7.7	Discussion	203
	References	205
8	Am I Going to Do This? Is It Me Who Did This?	209
	Prior and Retrospective Ownership States for Actions	209
8.1	Introduction	209
8.2	Neurological Background	211
8.3	A Temporal-Causal Network Model for Ownership	213
8.3.1	Conceptual Representation of the Temporal-Causal Network Model	213
8.3.2	Numerical Representation of the Temporal-Causal Network Model	215
8.4	Simulation of Example Scenarios	220
8.4.1	Normal Execution and Attribution of an Action	221
8.4.2	Vetoing a Prepared Action Due to Unsatisfactory Predicted Effect	222

8.4.3	Effects of Poor Prediction; Schizophrenia Case	224
8.4.4	Satisfactory Predicted Effects but Unsatisfactory Actual Effects	225
8.4.5	Mirroring Another Person	226
8.5	Relations to Neurological Findings	227
8.6	Discussion.	230
	References.	231
9	How Empathic Are You	235
	Displaying, Regulating, and Learning Adaptive Social Responses	235
9.1	Introduction	235
9.2	Neurological Background	237
9.2.1	Mirror Neurons	237
9.2.2	Control and Self-other Distinction	238
9.2.3	Emotion Integration	239
9.2.4	Enhanced Sensory Processing Sensitivity and Emotion Regulation	239
9.2.5	Empathic Responses	241
9.3	The Temporal-Causal Network Model	243
9.3.1	Conceptual Representation of the Model	243
9.3.2	Numerical Representation of the Temporal-Causal Network Model	247
9.4	Types of Social Response Patterns Shown	252
9.4.1	Overview of Basic Patterns	252
9.4.2	Oscillatory Patterns: Limit Cycle Behaviour	255
9.4.3	Comparison to Empirical Gaze Data	256
9.4.4	Interaction of Two Persons Displaying Regulation of Enhanced Sensory Sensitivity	257
9.5	Learning Social Responses by an Adaptive Temporal-Causal Network Model.	259
9.6	Example Simulations of Learning Processes	260
9.7	Discussion.	263
	References.	265
10	Are You with Me? Am I with You?	269
	Joint Decision Making Processes Involving Emotion-Related Valuing and Mutual Empathic Understanding.	269
10.1	Introduction	269
10.2	Mirroring, Internal Simulation and Emotion-Related Valuing.	270
10.3	The Temporal-Causal Network Model	272
10.3.1	Conceptual Representation of the Temporal-Causal Network Model	273
10.3.2	Numerical Representation of the Temporal-Causal Network Model	275

10.4	Simulation Results	278
10.5	Discussion.	281
	References.	282
11	Changing Yourself, Changing the Other, or Changing Your Connection	285
	Integrative Dynamics of States and Interactions in a Social Context	285
11.1	Introduction	285
11.2	Small World Networks and Random Networks.	286
	11.2.1 Small World Networks	288
	11.2.2 Random Networks	288
11.3	Distribution of Node Degrees and Scale-Free Networks	289
	11.3.1 Scale-Free Networks	289
	11.3.2 Identifying a Power Law	290
	11.3.3 Clusters and Bridges	292
11.4	Weak Ties, Strong Ties and Weighted Connections	292
11.5	Different Types of Dynamics in Networks Based on Social Interaction.	296
11.6	Social Contagion.	299
11.7	Adaptive Network Dynamics and the Homophily Principle	304
11.8	Adaptive Networks and the More Becomes More Principle	311
11.9	Adaptive Networks and Actual Interaction Over Time	313
11.10	Discussion.	317
	References.	318
Part IV	Analysis Methods for Temporal-Causal Network Models	
12	Where Is This Going	323
	Verification by Mathematical Analysis	323
12.1	Introduction	323
12.2	Verifying a Temporal-Causal Network Model by Mathematical Analysis.	324
12.3	Mathematical Analysis for Equilibrium States: An Example	330
12.4	Mathematical Analysis for Equilibrium States: Scaled Sum Combination Function.	333
12.5	Mathematical Analysis for Equilibrium States: Hebbian Learning	336
	12.5.1 Analysis of Increase, Decrease or Equilibrium for Hebbian Learning Without Extinction	337
	12.5.2 Analysis of Increase, Decrease or Equilibrium for Hebbian Learning with Extinction	338
	12.5.3 How Much Activation Is Needed to Let ω Increase?	340

12.6	Mathematical Analysis for Equilibrium States: Homophily Principle	341
12.7	Mathematical Analysis for Behaviour Ending up in a Limit Cycle Pattern.	343
12.8	Discussion.	347
	References.	348
13	What Is Happening.	349
	Identifying and Verifying Emergent Patterns.	349
13.1	Introduction	349
13.2	Dynamic Properties and Temporal-Causal Network Models	351
13.2.1	A Temporal-Causal Network Model Describing Local Dynamics and Dynamic Properties Describing Patterns Emerging in Overall Dynamics.	351
13.2.2	Identifying Emergent Dynamic Properties for a Given Model.	352
13.2.3	Identifying Dynamic Properties Initially as Requirements for a Model	353
13.3	Dynamic Properties Versus Real World Dynamics: Validation, Monitoring, and Analysis	354
13.3.1	Validating Dynamic Properties Against Actual Real World Processes	355
13.3.2	Validating Dynamic Properties Against Patterns Reported in Literature	356
13.3.3	Monitoring and Analysis of Real World Processes Using Dynamic Properties	356
13.4	Dynamic Properties Versus Model Dynamics: Verification and Personalization.	356
13.4.1	Testing, Focusing and Analysis of a Model by Verifying It Against Dynamic Properties	357
13.4.2	Personalizing Characteristics of a Model Based on Dynamic Properties	357
13.4.3	Validation of a Model Based on Validated Dynamic Properties	358
13.5	Conceptual Representations of Dynamic Properties	358
13.6	Numerical-Logical Representations of Dynamic Properties.	363
13.6.1	Numerical Representations of State Relations	364
13.6.2	Using Numerical Representations Within a Dynamic Property Expression	366
13.6.3	Numerical-Logical Representation of a Dynamic Property Expression.	368

13.7	Types of Dynamic Properties and Their Representations	371
13.7.1	Basic State Relation, Achievement, Grounding, Representation, Ordering and Monotonicity Properties.	371
13.7.2	Maintenance, Peak, Speed, Equilibrium and Limit Cycle Properties	375
13.7.3	State Comparison, Trace Comparison and Trace Selection Properties.	380
13.8	Examples of Dynamic Properties in Some Case Studies.	383
13.9	Automatic Checking of Dynamic Properties.	387
13.10	Discussion.	389
	References.	390
14	Who are You	393
	Identifying Characteristics of Persons, Their Networks and Other Contextual Aspects by Parameter Estimation and Validation	393
14.1	Introduction	393
14.2	Determining Characteristics and the Use of Requirements	395
14.2.1	The Parameters in a Temporal-Causal Network Model	395
14.2.2	Direct Measuring of Characteristics of a Situation.	396
14.2.3	Using Requirements to Find Characteristics of a Situation.	397
14.2.4	Using Error Measures for Requirements	398
14.3	Description of an Example Model	400
14.4	Parameter Tuning by Exhaustive Search.	403
14.5	Parameter Estimation by Gradient Descent.	406
14.6	Parameter Estimation by Random Gradient Descent.	410
14.7	Parameter Estimation by Simulated Annealing	412
14.8	Discussion.	417
	References.	418
 Part V Philosophical, Societal and Educational Perspectives		
15	We Don't Believe in Ghosts, Do We?	421
	What Is It that Drives Dynamics.	421
15.1	Introduction	421
15.2	Is Motion of Nonliving Entities Driven by Ghosts?	424
15.2.1	Zeno About Arrows that Are Moving and Unmoving	424
15.2.2	Adding Anticipatory State Properties to Describe a State: Potentialities.	427
15.3	Is Motion of Living Entities Driven by Ghosts?.	428
15.3.1	Mental States Driving Motion	428

15.3.2	Can ‘Things of the Soul’ Move Objects?	429
15.4	Explaining Changed States by Introducing Potentialities.	430
15.4.1	Potentialities and Their Actualisation as a General Perspective on Dynamics.	430
15.4.2	Derivatives as Potentialities for Variables in Dynamical Systems	431
15.4.3	What Kind of State Properties Are Potentialities? . . .	432
15.4.4	Summary of Assumptions Underlying Potentialities	433
15.5	Potentialities in Physics	433
15.6	What Kind of Property Is a Potentiality: Getting Rid of Ghosts?.	435
15.6.1	Why Velocities and Derivatives by Themselves Are not Genuine State Properties.	436
15.6.2	Ghost-like Properties or Temporal Relations Involving Genuine Properties?.	438
15.7	Potentialities for Causal Relations and Transition Systems . . .	440
15.7.1	Transition Systems and Causal Relations.	440
15.7.2	Potentialities for Transition Systems and Causal Relations	441
15.8	Realisers for Potentialities and the Role of Differential Equations	442
15.8.1	Realisers of Mental States in Philosophy of Mind	442
15.8.2	Realisers of Potentialities from a More General Perspective.	443
15.8.3	Realisers for Derivatives: First-Order Differential Equations.	444
15.9	How to Explain Changed Potentialities.	446
15.9.1	Introducing Higher-Order Potentialities: Potentialities for Potentialities	447
15.9.2	Higher-Order Potentialities in Cognitive Models . . .	448
15.9.3	Mathematical Formalisation of Higher-Order Potentialities in Calculus	448
15.9.4	How to Get Rid of an Infinite Chain of Higher Order Potentialities by Realisers	449
15.10	Changed Potentialities Due to Interaction	450
15.10.1	Exchange of Potentialities by Interaction.	450
15.10.2	The Role of Higher-Order Potentialities in the Exchange of Potentialities	452
15.10.3	Higher-Order Potentialities to Characterise Interaction in Physics.	453

15.11	Multiple Realisation of Potentialities.	455
15.12	State-Determined Systems and Potentialities.	457
15.13	Discussion.	459
	References.	461
16	Making Smart Applications Smarter	463
	Societal Applicability of Computational Models	463
16.1	Introduction	463
16.2	Multidisciplinarity: The Ingredients	465
16.3	Combining the Ingredients	465
16.4	Coupled Reflective Systems	467
16.5	Integrative Modeling.	468
16.6	Discussion.	470
	References.	471
17	Multidisciplinary Education	473
	Computational Modeling as the Core of a Multidisciplinary Curriculum	473
17.1	Introduction	473
17.2	Overall Structure of the Curriculum	475
17.3	Computational Modeling Stream.	477
17.4	The Human Sciences and Exact Sciences Streams	479
17.5	Integration and Projects	480
17.6	Evaluation and Discussion	480
	References.	483
 Part VI Network-Oriented Modeling: Discussion		
18	On the Use of Network-Oriented Modeling.	487
	A Discussion.	487
18.1	Introduction	487
18.2	Network-Oriented Modeling	487
18.3	Genericity of a Network-Oriented Modeling Approach.	488
18.4	Applicability of Network-Oriented Modeling	490
18.5	Finally	492
	References.	492
Index		495

Part I
Network-Oriented Modeling: Introduction

Chapter 1

Network-Oriented Modeling and Its Conceptual Foundations

An Introduction

Abstract To address complexity of modeling the world's processes, over the years in different scientific disciplines isolation and separation assumptions have been made, and in some disciplines they have turned out quite useful. They traditionally serve as a means to address the complexity of processes by some strong form of decomposition. It can be questioned whether such assumptions are adequate to address complexity of integrated human mental and social processes and their interactions. Are there better alternative strategies to address human complexity? This is discussed in this chapter, and it is pointed out that a Network-Oriented Modeling perspective can be considered an alternative way to address complexity, which is better suited for modeling human and social processes.

1.1 Introduction

To address complexity of modeling the world's processes, over the years different strategies have been used. From these strategies isolation and separation assumptions are quite common in all scientific disciplines and have often turned out very useful. They traditionally serve as means to address the complexity of processes by some strong form of decomposition. This also holds for classical disciplines such as Physics, where, for example, for mechanical modeling for building construction only forces from objects on earth are taken into account and not forces from all other objects in the universe, that still do have some effects as well. It is recognized that these distant effects from sun, moon, planets and other objects do exist, but it assumed that they can be neglected. For such cases within Physics such an isolation assumption may be a reasonable choice, but in how far is it equally reasonable to address complexity of human mental and social processes? Over the years within the Behavioural and Social Sciences also a number of assumptions have been made in the sense that some processes can be studied by considering them as separate or isolated phenomena. However, within these human-directed sciences serious debates or disputes have occurred time and time again on such a kind of assumptions.

They essentially have the form of arguments pro or con an assumption that some processes can be studied by considering them as separate or isolated phenomena. Examples of such separation assumptions to address human complexity concern:

- mind versus body
- cognition versus emotion
- individual processes versus collective processes
- non-adaptive processes versus adaptive processes
- earlier versus later: temporal separation

It can be questioned whether, for example, mind can be studied while ignoring body, or cognition while ignoring emotion, or sensory processing in isolation from action preparation. Or, put more general, in how far are these traditional means to address complexity by separation still applicable if the complexity of human mental and social processes has to be addressed? Do we need to break with such traditions to be able to make more substantial scientific progress in this area addressing human processes? And, not unimportant, are there adequate alternative strategies to address human complexity?

In this chapter, first in Sect. 1.2 the five separation assumptions mentioned above are discussed in some more detail. Next, in Sect. 1.3 it is discussed how as an alternative, interaction in networks can be used to address complexity. In Sect. 1.4 the development of a Network-Oriented Modeling perspective is discussed. Section 1.5 focuses on the need for a temporal dimension to address the dynamics, in particular to handle cyclic causal connections and realistic timing in human processes. In Sect. 1.6 the Network-Oriented Modeling approach based on temporal-causal networks is briefly pointed out, which is the modeling approach used in this book, and is discussed more extensively in Chap. 2. Section 1.7 discusses the scope of applicability of the approach. Finally, Sect. 1.8 provides an overview of the chapters in the book.

1.2 Addressing Human Complexity by Separation Assumptions

The position taken in this book is that indeed a number of the traditional separation and isolation habits followed in order to address human complexity have to be broken to achieve more progress in scientific development. Partly due to the strong development of Cognitive, Affective and Social Neuroscience, in recent years for many of the issues mentioned above, a perspective in which dynamics, interaction and integration are key elements has become more dominant: a perspective with interaction as a point of departure instead of separation. Given this background, for each of the separation issues listed above this will be discussed below in more detail. It will be pointed out how in many cases separation assumptions as mentioned lead to some types of discrepancies or paradoxes.

Mind versus Body

A first isolation assumption that has a long tradition is the assumption that the mind can be studied in separation from the body. There has been debate about this since long ago. Aristotle (350 BC) referred to properties of ‘mind and desire’ as the source of motion of a living being: he discusses how the occurrence of certain internal (mental) state properties (desires) within a living being entails or causes the occurrence of an action in the external world; see also Nussbaum (1978). Such internal state properties are sometimes called by him ‘things in the soul’, ‘states of character’, or ‘moral states’. In that time such ‘things’ were not considered part of the physical world but of a ghost-like world indicated in this case by ‘soul’. So, in this context the explanation that such a creature’s position gets changed is that there is a state of the soul driving it. This assumes a separation between the soul on the one hand, and the body within the physical world on the other hand. How such nonphysical states can affect physical states remains unanswered. Over time, within Philosophy of Mind this has been felt as a more and more pressing problem. The idea that mental states can cause actions in the physical world is called *mental causation* (e.g., Kim 1996, 1998). The problem with this is how exactly non-physical mental states can cause effects in the physical world, without any mechanism known for such an effect. Within Philosophy of Mind a solution for this has been proposed in the form of a tight relation between mental states and brain states. Then it is in fact not the mental state causing the action, but the corresponding (physical) brain state. Due to this the separation is not between the soul or mind, and the body, but between the brain and the body (Bickle 1998; Kim 1996, 1998).

However, this separation between brain and body also has been debated. More literature on this from a wider perspective can be found, for example, in Clark (1998), Lakoff and Johnson (1999), Wilson (2002). It is claimed that mind essentially is embodied: it cannot be isolated from the body. One specific case illustrating how brain and body intensely work together and form what is called an embodied mind is the causal path concerning feelings and emotional responses. A classical view is that, based on some sensory input, due to internal processing emotions are felt, and based on this they are expressed in some emotional response in the form of a body state, such as a face expression:

stimulus → sensory representation → felt emotion
→ preparation for a body state → expressed emotion in body state

However, James (1884) claimed a different order in the causal chain (see also Damasio 2010, pp. 114–116):

stimulus → sensory representation → preparation for a body state
 → expressed emotion in body state → sensed body state
 → representation of body state → felt emotion

The mechanisms of emotion and cognition appear to be intertwined at all stages of stimulus processing and their distinction can be difficult. (...) Adding the complexity of emotion to the study of cognition can be daunting, but investigations of the neural mechanisms underlying these behaviors can help clarify the structure and mechanisms (Phelps 2006, pp. 46–47).

Here it is recognized that an assumption on isolating cognition from emotion is not realistic, as far as the brain is concerned. Therefore models based on such an assumption cannot be biologically plausible and may simply be not valid. Moreover, it is also acknowledged that taking into account the intense interaction between emotion and cognition ‘can be daunting’; to avoid this problem was a main reason for the isolation assumption as a way to address complexity. However, Phelps (2006) also points at a way out of this: use knowledge about the underlying neural mechanisms. In the past when there was limited knowledge about the neural mechanisms this escape route was not available, and therefore the isolation assumption may have made sense, although the validity of the models based on that can be questioned. But, now Neuroscience has shown a strong development, this provides new ways to get rid of this isolation assumption. Similar claims about the intense interaction between emotion and cognition have been made by Pessoa (2008). In experimental contexts different types of effects of affective states on cognitive states have indeed been found; see, for example, Eich et al. (2000), Forgas et al. (2009), Winkielman et al. (2009), Frijda et al. (2000). Moreover, more specifically in the rapidly developing area of Cognitive Neuroscience (e.g., Purves et al. 2008; Gazzaniga 2009) knowledge has been contributed on mechanisms for the interaction and intertwining of affective and cognitive states and processes (for example, involving emotion, mood, beliefs or memory); see, for example, Dolan (2002), LaBar and Cabeza (2006), Pessoa (2008), Phelps (2006), Storbeck and Clore (2007).

Not only for desiring and believing the isolation assumption for cognition versus emotion is questioned, but also for rational decision making. Traditionally, rationality and emotions often have been considered each other’s enemies: decision making has often been considered as a rational cognitive process in which emotions can only play a disturbing role. In more recent times this has been questioned as well. For example, in Loewenstein and Lerner (2003, p. 619) it is pointed at the positive functions served by emotions:

Throughout recorded human intellectual history there has been active debate about the nature of the role of emotions or ‘passions’ in human behavior, with the dominant view being that passions are a negative force in human behavior (...). By contrast, some of the latest research has been characterized by a new appreciation of the positive functions served by emotions (Loewenstein and Lerner 2003, p. 619)

In particular, in decision making it may be questioned whether you can make an adequate decision without feeling good about it. Decisions with bad feelings associated to them may lack robustness. Many occasions may occur over time that trigger a temptation to change it into a decision with a better associated feeling. So, human experience in rational decisions and feelings about them is that they go or

should go in hand in hand and are not isolated. This indicates another paradox or discrepancy between the isolation assumption and how real life is experienced: emotions can be considered a vehicle for rationality (for more details, see Chap. 6). A brief sketch of the alternative perspective is as follows. Decision making usually considers a number of options for a choice to be made. Such a choice is often based on some form of valuing of the options. In this valuing process emotions come in: the predicted effect of some of the options relate to a more positive feeling than for other options. It has been found that such valuations relate to amygdala activations (see, e.g., Morrison and Salzman 2010; Murray 2007; Salzman and Fusi 2010). As valuing can be seen as a grounding for a decision, it turns out that an emotional type of grounding is involved. Bad decisions are those that are not solidly grounded by having a positive feeling about them. They may not last long, as any opportunity to get rid of them will be a temptation to reconsider the decisions. Recent neurological literature addressing this idea of emotional valuing and grounding of decisions relates the notion of value to the amygdala; e.g., Bechara et al. (2003), Bechara et al. (1999), Montague and Berns (2002), Janak and Tye (2015), Jenison et al. (2011), Morrison and Salzman (2010), Ousdal et al. (2014), Pessoa (2011), Rangel et al. (2008).

In Chap. 3 it is discussed how knowledge from Neuroscience can be used to find out how the integration of emotions and cognitive processes can be modeled, illustrated for a number of examples. In Chaps. 4 and 5, more specifically the role of emotions in generating dreams and learning during dreaming is discussed. In Chap. 6 the specific case of emotions as a basis for rational decision making is addressed in more detail.

Individual versus Collective

Yet another isolation assumption concerns the distinction between mental processes within an individual and social processes. The former are usually referred to the territory of Psychology, whereas the latter are referred to the territory of Social Science. The idea then is to study social processes as patterns emerging from interactions between individuals thereby abstracting from the processes within each of the individuals. This easily leads to some kind of paradoxes. For example, as persons in a group are autonomous individuals with their own neurological structures and patterns, carrying, for example, their own emotions, beliefs, desires and intentions, it would be reasonable to expect that it is very difficult or even impossible to achieve forms of sharedness and collectiveness. However, it can be observed that often groups develop coherent views and decisions, and, even more surprisingly, the group members seem to share a positive feeling about it. In recent years by developments in Neuroscience new light has been shed on this seeming paradox of individuality versus sharedness and collectiveness. This has led to the new discipline called Social Neuroscience; e.g., Cacioppo and Berntson (2005), Cacioppo et al. (2006), Decety and Cacioppo (2010), Decety and Ickes (2009), Harmon-Jones and Winkielman (2007). Two interrelated core concepts in this discipline are mirror neurons and internal simulation of another person's mental processes. Mirror neurons are neurons that not only have the function to prepare for

a certain action or body change, but are also activated upon observing somebody else who is performing this action or body change; e.g., Iacoboni (2008), Pineda (2009), Rizzolatti and Sinigaglia (2008). Internal simulation is internal mental processing that copies processes that may take place externally, for example, in mental processes in another individual; e.g., Damasio (1994, 1999), Gallese and Goldman (1998), Goldman (2006), Hesslow (1994, 2002, 2012). Mechanisms involving these core concepts have been described that provide an explanation of the emergence of sharedness and collectiveness from a biological perspective. This new perspective breaks the originally assumed separation between processes within individuals and processes of social interaction. This perspective is discussed in more detail in Chap. 7.

Adaptive versus Nonadaptive Processes

Another assumption that sometimes is debated is that mental and social processes are modeled as if they are not adaptive. In reality processes usually have adaptive elements incorporated, but often these elements are neglected and sometimes studied as separate phenomena. One example in a social context is the following. Often a contagion principle based on social interaction is studied, describing how connected states affect each other by these interactions, whereas the interactions themselves are assumed not to change over time (for example, qua strength, frequency or intensity). But in reality the interactions also change, for example based on what is called the *homophily principle*: the more you are alike, the more you like (each other); for example, see Byrne (1986), McPherson et al. (2001), Mislove et al. (2010). Another way of formulating this principle is: birds of a feather flock together. It can often be observed that persons that have close relationships or friendships are alike in some respects. For example, they go to the same clubs, watch the same movies or TV programs, take the same drinks, have the same opinions, vote for the same or similar parties. Such observations might be considered support for the contagion principle: they were together and due to that they affected each other's states by social contagion, and therefore they became alike. However, also a different explanation is possible based on the homophily principle: in the past they already were alike before meeting each other, and due to this they were attracted to each other. So, the cyclic relation between *the states of the members* and *the strength of their connection* leads to two possible causal explanations of being alike and being connected:

being connected $\rightarrow$ being alike	(contagion principle)
being alike $\rightarrow$ being connected	(homophily principle)

Such circular causal relations make it difficult to determine what came first. It may be a state just emerging from a cyclic process without a single cause. For more discussion on this issue, for example, see Aral et al. (2009), Shalizi and Thomas (2011), Steglich et al. (2010), Mundt et al. (2012). This phenomenon will be addressed in more detail in Chap. 11.

As another example illustrating how adaptivity occurs fully integrated with the other processes, the function of dreaming is discussed. From a naïve perspective, dreaming might be considered as just playing some movie, thereby triggering some emotions, and that's all. But in recent research, the idea has become common that dreaming is a form of internal simulation of real-life-like processes serving as training in order to learn or adapt certain capabilities. Dreaming makes use of memory elements for sensory representations (mental images) and their associated emotions (learnt in the past) to generate 'virtual simulations'; e.g., Levin and Nielsen (2007, pp. 499–500). Taking into account fear emotions that often play an important role in dreams, strengthening of regulation of such emotions is considered an important purpose of dreaming; see, for example, Levin and Nielsen (2007), Walker and van der Helm (2009), van der Helm et al. (2011), Gujar et al. (2011), Deliens et al. (2014), Pace-Schott et al. (2015), Sotres-Bayon et al. (2004). To this end in dreams adequate exercising material is needed: sensory representations of emotion-loaden situations are activated, built on memory elements suitable for high levels of arousal. The basis of what is called 'fear extinction learning' is that emotion regulation mechanisms are available which are adaptive: they are strengthened over time when they are intensively used. Fear extinction learning as an expression may sound a bit paradoxal; it is *not* a form of unlearning or extinction of acquired fear associations, but it is additional *learning of fear inhibition* connections in order to counterbalance the fear associations which themselves remain intact (e.g., Levin and Nielsen 2007, p. 507). Such a strengthening of connections can be described by a Hebbian learning principle (Hebb 1949); see also Chap. 2, Sect. 2.10. The processes of dreaming and the adaptive elements involved in it are addressed in Chap. 5.

Earlier versus Later: Temporal Separation

Another traditionally made separation assumption is that processes in the brain are separated in time. For example, sensing, sensory processing, preparation for action and action execution are assumed to occur in linearly ordered sequential processes:

sensing – sensory processing – preparing for action – executing action

For the case of emotions it was already discussed that such linear temporal patterns are not applicable. But also more in general it can be argued that such linear patterns are too much of a simplification, as in reality such processes occur simultaneously, in parallel; often a form of internal simulation takes place, as put forward, among others, by Hesslow (1994, 2002, 2012), Damasio (1994, 1999), Goldman (2006), Barsalou (2009), Marques and Holland (2009), Pezzulo et al. (2013). The general idea of internal simulation that was also mentioned above in the specific context of emotions and bodily processes, is that sensory representation states are activated (e.g., mental images), which in response trigger associated preparation states for actions, which, by prediction links, in turn activate other sensory representation states for the predicted effects of the prepared actions:

sensory representation states → preparation states → sensory representation states

The latter states represent the effects of the prepared actions or bodily changes, without actually having executed them. Being inherently cyclic, the simulation process can go on indefinitely, as the latter sensory representations can again trigger preparations for actions, and so on, and everything simultaneously, in parallel, as in the world no process is freezing to wait for another process to finish first. Internal simulation has been used, for example, to describe (imagined) processes in the external world, e.g., prediction of effects of own actions (Becker and Fuchs 1985), or processes in another person's mind, e.g., emotion recognition or mindreading (Goldman 2006) or (as discussed above) processes in a person's own body by as-if body loops (Damasio 1994). This breaks with the tradition that there is a temporal separation of processes such as sensing—sensory processing—preparing for action—executing action. In many of the chapters this is illustrated.

1.3 Addressing Complexity by Interaction in Networks Instead of by Separation

The separation assumptions to address complexity as discussed in Sect. 1.2 are strongly debated, as they all come with shortcomings. In this section it is discussed that in fact the problem is not so much in the specific separation assumptions, but in the general idea of separation itself. In social contexts it is clear that the intense interaction between persons based on their mutual and often interrelated cyclic relationships, makes them not very well suitable for any separation assumptions: all these interactions take place all the time, simultaneously, in parallel. And this does not only apply to social processes but also to individual mental processes, as will be discussed in some more detail here.

In the domain of Neuroscience the structures and mechanisms found suggest that many parts in the brain are connected by connections that often are part of cyclic paths, and such cycles are assumed to play an important role in many mental processes (e.g., Bell 1999; Crick and Koch 1998; Potter 2007). As an example also put forward above, there is a growing awareness, fed by findings in Neuroscience that emotions play an important mediating role in most human processes, and this role often provides a constructive contribution, and not a disturbing contribution as was sometimes assumed. Usually mental states trigger emotions and these emotions in turn affect these and other mental states. It turns out that to address this type of circular effects, different views on causality and modeling are required, compared to the traditional views in modeling of mental processes. For example, Scherer (2009) states:

What is the role of causality in the mechanisms suggested here? Because of the constant recursivity of the process, the widespread notion of linear causality (a single cause for a single effect) cannot be applied to these mechanisms. Appraisal is a process with constantly changing results over very short periods of time and, in turn, constantly changing driving effects on subsystem synchronization (and, consequently, on the type of emotion). (...) Thus, as is generally the case in self-organizing systems, there is no simple, unidirectional sense of causality (see also Lewis 1996). (Scherer 2009, p. 3470)

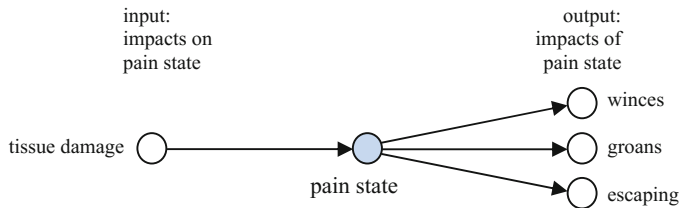


Fig. 1.1 Pain state with some of its causal relations

Also in the domain of Philosophy of Mind this issue of cyclic causal connections is recognized, for example, by Kim (1996). The idea is that a mental state is characterized by the way it mediates between the input it receives from other states and the output it provides to other states; this is also called the *functional* or *causal role* of the mental state. The idea is that each mental state is characterized by its causal role. For example, as a simplified example on the input side a mental state of being in pain is typically caused by tissue damage and in turn on the output side it typically causes wincing, groans and escape behavior (Kim 1996, p. 104); see Fig. 1.1. So, in this perspective the question what exactly is pain can be answered as the state that forms a causal bridge (or causally mediates) from tissue damage to wincing, groans, and escape behavior. Kim describes the overall picture as follows:

Mental events are conceived as nodes in a complex causal network that engages in causal transactions with the outside world by receiving sensory inputs and emitting behavioral outputs (Kim 1996, p. 104).

As input not only sensory input can play a role but also input from other mental states such as in the pain example ‘being alert’. Similarly, as output not only behavioral output can play a role but also other mental states can be affected, such as in the pain example feeling distress and a desire to be relieved of it. Within Philosophy of Mind this is often considered challenging:

But this seems to involve us in a regress or circularity: to explain what a given mental state is, we need to refer to other mental states, and explaining these can only be expected to require reference to further mental states, on so on – a process that can go on in an unending regress, or loop back in a circle (Kim 1996, pp. 104–105).

In Fig. 1.2 an example of such a cyclic causal path is depicted. Here mental state S_1 has a causal impact on mental state S_2 , but one of the states on which S_2 has an effect, in turn affects one of the input states for S_1 .

This view from Philosophy of Mind is another indication that a modeling approach will have to address causal relations with cycles well. To obtain an adequate understanding of such cycles and their dynamics and timing it is inevitable to take into account the temporal dimension of the dynamics of the processes effectuated by the causal relations. In principle, this situation makes that an endless cyclic process over time emerges, which in principle works simultaneously, in parallel, and in interaction with other processes. In such a graph at each point in

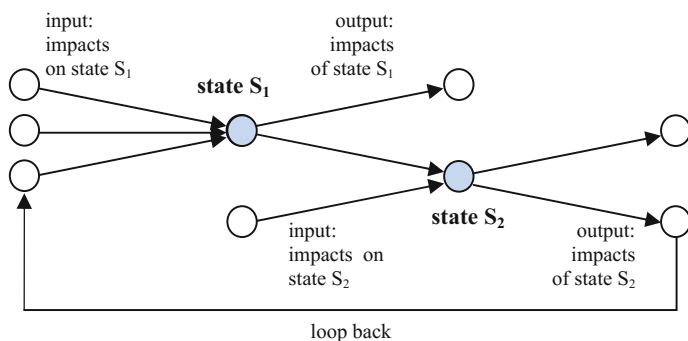


Fig. 1.2 Mental states with their causal relations conceived as nodes in a complex, often cyclic causal network (see also Kim 1996, p. 104)

time activity takes place in every state simultaneously (it is not that one state waits for the other). The notion of state at some point in time used here refers to a specific part or aspect of the overall state of a model at this point in time. Such an overall state can include, for example, at the same time a ‘being in pain’ state, a ‘desire to get relieved’ state, and an ‘intention to escape’ state. The overall state at some point in time is the collection of all states at that point in time. All the time the processes in the brain occur in parallel, in principle involving all specific states within the overall state, mostly in an unconscious manner. In this sense the brain is not different from any other part of the universe where everywhere processes take place simultaneously, in parallel. During all this parallel processing, any change in state S_1 in principle will lead to a change in state S_2 , which in turn will lead to another change in state S_1 , which leads to another change in state S_2 , and so on and on. The state changes in such a process may become smaller and smaller over time, and the cyclic process eventually may converge to an equilibrium state in which no further changes occur anymore; but also other patterns are possible, such as limit cycles in which the changes eventually end up in a regular, periodic pattern of changes (see also Chap. 12).

In the sense described above, mental processes can show patterns similar to patterns occurring in social interactions, where cycles of connections are natural and quite common. An example from the context of modeling social systems or societies can be found in (Naudé et al. 2008):

The paper outlines the challenges of modeling and assessing spatially complex human-ecosystem interactions, and the need to simultaneously consider rural-urban and rich-poor interactions. The context for exploring these challenges is South Africa, which has such stark poor-rich and associated rural-urban and other spatial disparities, that it is often described as a microcosm of the global division between developed and developing countries. Instead of rigid rural-urban dichotomies and other absolute, “container” views of space, there is a need to recognise spatial overlaps and complexities such as the pervasiveness of so-called translocal livelihood systems. Accordingly, much more *relational, network-oriented modeling approaches* are needed (Naudé et al. 2008, p. 1).

Also here it is claimed that separation in the form of what they call ‘container’ views of space falls short in addressing the complexities involved, and as an alternative a Network-Oriented Modeling approach is suggested to address human social complexity. Similar claims are made from the area of organization modeling by Elzas (1985):

The study of the process-type of organization, which still - at this moment because its relative novelty - requires modeling to evaluate, can benefit from certain network-oriented modeling formalisms because of the very nature of the organizational concept. (...) in addressing (...) the specific coordinating problems of the adaptively interrelated distributed-action organizational units as they are found in process-based organizational models (Elzas 1985, p. 162).

So, both from the area of the analysis of mental processes and from the area of analysis of social processes, the notion of network is suggested as a basis. In next section the notion of Network-Oriented Modeling is discussed in some more detail.

1.4 Network-Oriented Modeling

This chapter started in Sect. 1.2 by some reflection on traditional means to address complexity by assuming separation and isolation of processes, and the shortcomings, discrepancies and paradoxes entailed by these assumptions. In Sect. 1.3 the circular or cyclic, interactive and distributed character of many processes (involving interacting sub-processes running simultaneously, in parallel) was identified as an important challenge to be addressed, and it was recognized that a perspective based on interactions in networks is more suitable for this. In this section a Network-Oriented Modeling perspective is proposed as an alternative way to address complexity. This perspective takes the concept of network and the interactions within a network as a basis for conceptualization and structuring of any complex processes. Network-Oriented Modeling is *not* considered here as modeling *of* (given) networks, but modeling any (complex) processes *by* networks. It is useful to keep in mind that the concept network is a just a mental concept and this is used as a conceptual structuring tool to conceptualize any processes that exist in reality.

The concept of network is easy to visualize on paper, on a screen or mentally and as such provides a good support for intuition behind a model. Moreover, as the Network-Oriented Modeling approach presented here (see Sect. 1.6) also incorporates a temporal dimension enabling interpretation of connections as temporal-causal connections, the mental concept of network also provides support for the intuition behind the dynamics of the modeled processes.

The scientific area of networks has already a longer tradition within different disciplines of more than 60 years. But it has developed further and within many other disciplines, such as Biology, Neuroscience, Mathematics, Physics, Economics, Informatics or Computer Science, Artificial Intelligence, and Web Science; see, for example Boccaletti et al. (2006), Valente (2010), Giles (2012).

These developments already show how processes in quite different domains can be conceptualized as networks. Historically the use of the concept network in different domains can be traced back roughly to the years 1930–1950, or even earlier, for studying processes such as:

- brain processes in Neuroscience by neural networks; e.g. McCulloch and Pitts (1943), Rosenblatt (1958)
- metabolic processes in Cell Biology by metabolic networks; e.g., Ouellet and Benson (1951), Westerhoff et al. (1984)
- social interactions within Social Science by social networks; e.g., Bott (1957), Aldous and Straus (1966)
- processes in Human Physiology; e.g., Huber (1941), Wiener and Rosenblueth (1946)
- processes in engineering in Physics; e.g., Hubbard (1931), Bode (1945)
- processes in engineering in Chemistry; e.g., Treloar (1943), Flory (1944)

Within such literature often graphical representations of networks are used as an important means of presentation. After getting accustomed to such conceptualizations as networks of processes that exist in the real world, a belief may occur that these networks actually exist in reality (as neural networks, or as computer networks, or as social networks, for example), and *modeling by networks* happens sometimes to be phrased alternatively as *modeling networks*. However, it still has to be kept in mind that the concept ‘network’ is a mental concept used as a tool to conceptualize any type of processes. To make this distinction more clear linguistically, the phrase Network-Oriented Modeling is used as indication for modeling by networks. Within this book the preferred use of the word ‘network’ is to indicate a model or conceptualization of some process, not to indicate the process in the real world itself. For example, social media such as Facebook, Twitter, WhatsApp, Instagram,... do not form or create social networks in reality, but they create social interactions in reality that can be described (conceptualized, modeled) by (social) networks or by network models.

Network-Oriented Modeling offers a conceptual tool to model complex processes in a structured, intuitive and easily visualizable manner, but the approach described here also incorporates the dynamics of the processes in these models. Using this approach, different parts of a process can be distinguished, but in contrast to the separation and isolation strategy to address complexity, a network-oriented approach does not separate or isolate these parts, but emphasizes and explicitly models the way how they are connected and interact. Moreover, by adding a temporal dimension to incorporate a dynamic perspective, it is explicitly modeled how they can have intense and circular causal interaction, and how the timing of the processes is. As intense interaction in network models as a way of modeling requires a dynamic, temporal perspective, this will be discussed next.

1.5 The Dynamic Computational Modeling Perspective

The challenge to cope with a dynamical and cyclic picture of both mental processes and social interaction processes, imposes certain requirements on a modeling approach. The modeling approach has to be able to handle time and dynamics well. For example, in (van Gelder and Port 1995) the symbolic computational perspective is criticized as being not able to address the time-context of cognitive processes in an adequate manner. In contrast they propose a perspective in which cognition is considered as dynamics:

The alternative, then, is the *dynamical* approach. Its core is the application of the mathematical tools of dynamics to the study of cognition. (...) But the dynamical approach is more than just powerful tools; like the computational approach it is a worldview. The cognitive system is not a computer, it is a dynamical system. (...) The cognitive system is not a discrete sequential manipulation of static representational structures; rather, it is a structure of mutually and simultaneously influencing *change* (van Gelder and Port 1995, p. 3).

They compare the dynamical perspective to the symbolic computational perspective as described by Newell and Simon's (1976) *Physical Symbol System Hypothesis*:

According to this hypothesis, natural cognitive systems are intelligent by virtue of being physical symbol systems of the right kind. At this same level of generality, dynamicists can be seen as embracing the *Dynamical Hypothesis*: Natural cognitive systems are dynamical systems, and are best understood from the perspective of dynamics. Like its computational counterpart, the Dynamical Hypothesis forms a general framework within which detailed theories of particular aspects of cognition can be constructed (van Gelder and Port 1995, p. 5).

It has taken a number of years before the dynamical perspective was adopted more substantially in practical cognitive and neuroscientific modeling work; see for example:

Although the idea of applying dynamical systems theory to the study of neural and cognitive mechanisms has been around for at least two decades (Beer 2000; Kelso 1995; Thelen and Smith 1994; van Gelder 1998), the dynamical systems approach has only recently begun to figure prominently in neuroscience (...) (Schurger and Uithol 2015).

The notion of *state-determined system*, adopted from Ashby (1960) is taken by van Gelder and Port (1995) as a definition of what a dynamical system is:

A system is state-determined only when its current state always determines a unique future behaviour. Three features of such systems are worth noting.

First, in such systems, the future behaviour cannot depend in any way on whatever states the system might have been in *before* the current state. In other words, past history is irrelevant (or at least, past history only makes a difference insofar as it has left an effect on the current state).

Second, the fact that the current state determines future behaviour implies the existence of some *rule of evolution* describing the behaviour of the system as a function of its current state. For systems we wish to understand we always hope that this rule can be specified in

some reasonable succinct and useful fashion. One source of constant inspiration, of course, has been Newton's formulation of the laws governing the solar system.

Third, the fact that future behaviours are uniquely determined means that state space sequences can never fork (van Gelder and Port 1995), p. 6.

Ashby (1960) emphasizes the importance of the identification of state-determined systems in a wide variety of scientific domains; for more details on this notion, see Chap. 2, Sect. 2.9. This perspective on mental systems as state-determined dynamical systems put forward by Ashby (1960) and van Gelder and Port (1995) can be viewed as a further extension of the world view for the universe as developed much earlier, for example, by Descartes. As also discussed in Treur (2007, Sects. 2.1 and 2.2, pp. 58–59), Descartes (1634) introduced a perspective on the world that sometimes is called the *clockwork universe*. This perspective claims that with sufficiently precise understanding of the world's dynamics at some starting time, the future can be predicted just by applying a set of 'laws of nature'. He first describes how at some starting time matter came into existence in a diversity of form, size, and motion. From that time on, dynamics continues according to these laws of nature.

From the first instant that they are created, He makes some begin to move in one direction and others in another, some faster and others slower (or indeed, if you wish, not at all); thereafter, He makes them continue their motion according to the ordinary laws of nature. For God has so wondrously established these laws that, even if we suppose that He creates nothing more than what I have said, and even if He does not impose any order or proportion on it but makes of it the most confused and most disordered chaos that the poets could describe, the laws are sufficient to make the parts of that chaos untangle themselves and arrange themselves in such right order that they will have the form of a most perfect world, in which one will be able to see not only light, but also all the other things, both general and particular, that appear in this true world (Descartes 1634, Chap. 6: Description of a New World, and on the Qualities of the Matter of Which it is Composed).

Descartes emphasizes that after such a starting time nothing (even no God) except the laws of nature determines the world's dynamics:

Know, then, first that by "nature" I do not here mean some deity or other sort of imaginary power. Rather, I use that word to signify matter itself, insofar as I consider it taken together with all the qualities that I have attributed to it, and under the condition that God continues to preserve it in the same way that He created it. For from that alone (i.e., that He continues thus to preserve it) it follows of necessity that there may be many changes in its parts that cannot, it seems to me, be properly attributed to the action of God (because that action does not change) and hence are to be attributed to nature. The rules according to which these changes take place I call the "laws of nature" (Descartes 1634, Chap. 7: On the Laws of Nature of this New World).

This view on the world's dynamics is often compared to a clockwork. The view assumes that systematic relationships (laws of nature) are possible between world states over time, in the sense that (properties of) past world states entail (properties of) future world states. The clockwork universe view has been developed further by Newton, Leibniz, Laplace and others. The following quotation taken from Laplace (1825) sketches how an intellect could be able to determine (by means of 'a single

formula') future world states from a present world state, that by itself is the effect of past world states:

We may regard the present state of the universe as the effect of its past and the cause of its future. An intellect which at any given moment knew all of the forces that animate nature and the mutual positions of the beings that compose it, if this intellect were vast enough to submit the data to analysis, could condense into a single formula the movement of the greatest bodies of the universe and that of the lightest atom; for such an intellect nothing could be uncertain and the future just like the past would be present before its eyes (Laplace 1825).

The worldview of Descartes and others described above in principle focuses on the physical universe. As such it applies to all physical and also biological processes in the universe, for example, those in the brain. The dynamical perspective on cognition put forward by Ashby (1960) and van Gelder and Port (1995) can be viewed as an extension of the above worldview from the physical world to the mental world. As this dynamical worldview already is assumed to apply to the physical processes in the brain, it is an advantage that also assuming such a worldview for mental processes will make it easier to relate mental and neural processes, as discussed in Chap. 2, Sect. 2.3.

1.6 Network-Oriented Modeling Based on Temporal-Causal Networks

As discussed above, both internal mental processing and social processing due to social interactions often involve multiple cyclic processes and adaptive elements. This has implications for the type of modeling approach to be used. Within Network-Oriented Modeling, the network models considered have to integrate such cycles, and also allow adaptive processes by which individuals can change their connections. To model such dynamics, a dynamical modeling perspective is needed that can handle such combinations of cycles and the adaptation of connections over time. Therefore, within the Network-Oriented Modeling approach as discussed here, the dynamic perspective has to be incorporated as well: a temporal dimension is indispensable. This is what is achieved in the Network-Oriented Modeling approach based on temporal-causal networks described in Chap. 2; see also Treur (2016).

The Network-Oriented Modeling approach based on temporal-causal networks is a generic and declarative dynamic AI modeling approach based on networks of causal relations (e.g., Kuipers and Kassirer 1983, Kuipers 1984, Pearl 2000), that incorporates a continuous time dimension to model dynamics. As discussed above, this temporal dimension enables causal reasoning and simulation for cyclic causal graphs or networks that usually inherently contain cycles, such as networks modeling mental or brain processes, or social interaction processes, and also the timing of such processes. States in such a network are characterised by the connections they have to other states, comparable to the way in which in Philosophy of Mind mental states are characterised by their causal roles, as discussed in Sect. 1.3. Moreover, adaptive elements can be fully integrated. The modeling approach can

incorporate ingredients from different modeling approaches, for example, ingredients that are sometimes used in specific types of (continuous time, recurrent) neural network models, and ingredients that are sometimes used in probabilistic or possibilistic modeling. It is more generic than such methods in the sense that a much wider variety of modeling elements are provided, enabling the modeling of many types of dynamical systems, as described in (Chap. 2, Sect. 2.9).

As discussed in detail in Chap. 2 and (Treur 2016) temporal-causal network models can be represented at two levels: by a conceptual representation and by a numerical representation. These model representations can be used not only to display interesting graphical network pictures, but also for numerical simulation. Furthermore, they can be analyzed mathematically and validated by comparing their simulation results to empirical data. Moreover, they usually include a number of parameters for domain, person, or social context-specific characteristics. To estimate values for such parameters, a number of parameter tuning methods are available.

A conceptual representation of a temporal-causal network model in the first place involves representing in a declarative manner states and connections between them that represent (causal) impacts of states on each other, as assumed to hold for the application domain addressed. The states are assumed to have (activation) levels that vary over time. What else is needed to describe processes in which causal relations play their role? In reality not all causal relations are equally strong, so some notion of *strength of a connection* is needed. Furthermore, when more than one causal relation affects a state, in which manner do these causal effects combine? So, some way to *aggregate multiple causal impacts* on a state is needed. Moreover, not every state has the same extent of flexibility; some states may be able to change fast, and other states may be more rigid and may change more slowly. Therefore, a notion of *speed of change* of a state is used for timing of processes. These three notions are covered by elements in the Network-Oriented Modeling approach based on temporal-causal networks, and are part of a conceptual representation of a temporal-causal network model:

- **Strength of a connection** $\omega_{X,Y}$

Each connection from a state X to a state Y has a *connection weight value* $\omega_{X,Y}$ representing the strength of the connection, often between 0 and 1, but sometimes also below 0 (negative effect) or above 1.

- **Combining multiple impacts on a state** $c_Y(..)$

For each state (a reference to) a *combination function* $c_Y(..)$ is chosen to combine the causal impacts of other states on state Y .

- **Speed of change of a state** η_Y

For each state Y a *speed factor* η_Y is used to represent how fast a state is changing upon causal impact.

Combination functions in general are similar to the functions used in a static manner in the (deterministic) Structural Causal Model perspective described, for

example, in Wright (1921), Pearl (2000), Mooij et al. (2013), but in the Network-Oriented Modeling approach described here they are used in a dynamic manner, as will be pointed out below briefly, and in more detail in Chap. 2.

Combination functions can have different forms. How exactly does one impact on a given state add to another impact on the same state? In other words, what types of combination functions can be considered? The more general issue of how to combine multiple impacts or multiple sources of knowledge occurs in various forms in different areas, such as the areas addressing imperfect reasoning or reasoning with uncertainty or vagueness. For example, in a probabilistic setting, for modeling multiple causal impacts on a state often independence of these impacts is assumed, and a product rule is used for the combined effect; e.g., Dubois and Prade (2002). In practical applications, this assumption is often questionable or difficult to validate. In the areas addressing modeling of uncertainty also other combination rules are used, for example, in possibilistic approaches minimum- or maximum-based combination rules are used; e.g., Dubois and Prade (2002). In another different area, addressing modeling based on neural networks yet another way of combining effects is used often. In that area, for combination of the impacts of multiple neurons on a given neuron usually a logistic sum function is used: adding the multiple impacts and then applying a logistic function; e.g., Grossberg (1969), Hirsch (1989), Hopfield (1982, 1984), Beer (1995).

So, there are many different approaches possible to address the issue of combining multiple impacts. The applicability of a specific combination rule for this may depend much on the type of application addressed, and even on the type of states within an application. Therefore the Network-Oriented Modeling approach based on temporal-causal networks incorporates for each state, as a kind of parameter, a way to specify how multiple causal impacts on this state are aggregated. For this aggregation a number of standard combination functions are made available as options and a number of desirable properties of such combination functions have been identified (see Chap. 2, Sects. 2.6 and 2.7), some of which are shown in Table 1.1.

These options cover elements from different existing approaches, varying from approaches considered for reasoning with uncertainty, probability, possibility or vagueness, to approaches based on recurrent neural networks; e.g., Dubois et al.

Table 1.1 Overview of some standard combination functions $c(V_1, \dots, V_k)$

Name	Description	Formula $c(V_1, \dots, V_k) =$
ssum_λ(..)	Scaled sum	$(V_1 \times \dots \times V_k)/\lambda$ with $\lambda > 0$
product(..) cproduct(..)	Product Complement product	$V_1 \times \dots \times V_k$ $1 - (1 - V_1) \times \dots \times (1 - V_k)$
min(..) max(..)	Minimal value Maximal value	$\min(V_1, \dots, V_k)$ $\max(V_1, \dots, V_k)$
slogistic_{σ,τ}(..)	Simple logistic sum	$1/(1 + e^{-\sigma(V_1 + \dots + V_k - \tau)})$ with $\sigma, \tau \geq 0$
alogistic_{σ,τ}(..)	Advanced logistic sum	$[(1/(1 + e^{-\sigma(V_1 + \dots + V_k - \tau)}) - (1/(1 + e^{\sigma\tau}))](1 + e^{-\sigma\tau})$ with $\sigma, \tau \geq 0$

(1991), Dubois and Prade (2002), Giangiacomo (2001), Zadeh (1978), Grossberg (1969), Hirsch (1989), Hopfield (1982, 1984), Beer (1995). In addition, there is still the option to specify any other (non-standard) combination function, preferably taking into account the desired properties.

The above three concepts (connection weight, speed factor, combination function) can be considered as parameters representing characteristics in a network model. In a non-adaptive network model these parameters are fixed over time. But to model processes by *adaptive networks*, not only the state levels, but also these parameters can change over time. For example, the connection weights can change over time to model evolving connections in network models.

A conceptual representation of a temporal-causal network model can be transformed in a systematic or even automated manner into a numerical representation of the model as follows (Treur 2016):

- at each time point t each state Y in the model has a real number value in the interval $[0, 1]$, denoted by $Y(t)$
- at each time point t each state X connected to state Y has an impact on Y defined as **impact** $_{X,Y}(t) = \omega_{X,Y} X(t)$ where $\omega_{X,Y}$ is the weight of the connection from X to Y
- The *aggregated impact* of multiple states X_i on Y at t is determined using a *combination function* $\mathbf{c}_Y(\cdot)$:

$$\begin{aligned} \mathbf{aggimpact}_Y(t) &= \mathbf{c}_Y(\mathbf{impact}_{X_1,Y}(t), \dots, \mathbf{impact}_{X_k,Y}(t)) \\ &= \mathbf{c}_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) \end{aligned}$$

where X_i are the states with connections to state Y

- The effect of **aggimpact** $_Y(t)$ on Y is exerted over time gradually, depending on speed factor η_Y :

$$\begin{aligned} Y(t + \Delta t) &= Y(t) + \eta_Y[\mathbf{aggimpact}_Y(t) - Y(t)]\Delta t \\ \text{or } dY(t)/dt &= \eta_Y[\mathbf{aggimpact}_Y(t) - Y(t)] \end{aligned}$$

- Thus, the following *difference* and *differential equation* for Y are obtained:

$$\begin{aligned} Y(t + \Delta t) &= Y(t) + \eta_Y[\mathbf{c}_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) - Y(t)]\Delta t \\ dY(t)/dt &= \eta_Y[\mathbf{c}_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) - Y(t)] \end{aligned}$$

For modeling processes as adaptive networks, some of parameters (such as connection weights) are handled in a similar manner, as if they are states. For more detailed explanation, see Chap. 2, Sect. 2.10.

Summarizing, as will be discussed in more detail in Chap. 2, the Network-Oriented Modeling approach based on temporal-causal networks described here provides a complex systems modeling approach that enables a modeler to design conceptual model representations in the form of networks described as cyclic graphs (or connection matrices), which can be systematically transformed into

executable numerical representations that can be used to perform simulation experiments. The modeling approach makes it easy to take into account on the one hand theories and findings from any domain from, for example, biological, psychological, neurological or social sciences, as such theories and findings are often formulated in terms of causal relations. This applies, among others, to mental processes based on complex brain processes, which, for example, often involve dynamics based on interrelating and adaptive cycles, but equally well it applies to social interaction processes and their adaptive dynamics. This enables to address complex adaptive phenomena such as the integration of emotions within all kinds of cognitive processes, of internal simulation and mirroring of mental processes of others, and dynamic social interaction patterns.

1.7 Scope of Applicability and Achievements

Concerning the scope of applicability, it has been shown (see Chap. 2, Sect. 2.9) that any smooth continuous state-determined system (any dynamical system described as a state-determined system or by a set of first order differential equations) can also be modeled by temporal-causal networks, by choosing suitable parameters such as connection weights, speed factors and combination functions. In this sense it is as general as modeling approaches put forward, for example, in Ashby (1960), Forrester (1973, 1987), Thelen and Smith (1994), Port and van Gelder (1995), van Gelder and Port (1995), Beer (1995), Kelso (1995), van Gelder (1998), and approaches such as described, for example in Grossberg (1969), Hopfield (1982, 1984), Hirsch (1989), Funahashi and Nakamura (1993).

To facilitate applications, dedicated software is available supporting the design of models in a conceptual manner, automatically transforming them into an executable format and performing simulation experiments. A variety of example models that have been designed illustrates the applicability of the approach in more detail, for example, as shown in a number of chapters in this book (see also Chap. 18, Sect. 18.4).

The topics addressed have a number of possible applications. An example of such an application is to analyse the spread of a healthy or unhealthy lifestyle in society. Another example is to analyse crowd behaviour in emergency situations. A wider area of application addresses socio-technical systems that consist of humans and devices, such as smartphones, and use of social media. For such mixed groups, in addition to analysis of what patterns may emerge, also for the support side the design of these devices and media can be an important aim, in order to create a situation that the right types of patterns emerge. This may concern, for example, safe evacuation in an emergency situation or strengthening development of a healthy lifestyle. Other application areas may address, for example, support and mediation in collective decision making and avoiding or resolving conflicts that may develop.

1.8 Overview of the Book

The book is composed of six parts:

- I. Network-Oriented Modeling: an Introduction
- II. Emotions all the Way
- III. Yourself and the Others
- IV. Analysis Methods for Temporal-Causal Network Models
- V. Philosophical, Societal and Educational Perspectives

For each part the chapters are briefly discussed here.

Part I Network-Oriented Modeling: An Introduction

This part is the introduction to the book, both conceptually and in a more technical sense. It consists of the current introduction Chap. 1, and a next Chap. 2 in which the Network-Oriented Modeling approach based on temporal-causal networks is introduced in detail.

Part II Emotions All the Way

In Part II a number of processes and models are discussed that address individuals and the way in which emotions are integrated in an interactive manner in practically all mental processes.

In Chap. 3 it is discussed how within Cognitive, Affective and Social Neuroscience more and more mechanisms have been found that suggest how emotions interact in a bidirectional manner with many other mental processes and behaviour. Based on this, in this chapter an overview of neurologically inspired temporal-causal network models for the dynamics and interaction for emotions is discussed. Thus an integrative perspective is obtained that can be used to describe, for example, how emotions interact with beliefs, experiences, decision making, and emotions of others, and also how emotions can be regulated. It is pointed out how integrated temporal-causal network models of such mental processes incorporating emotions can be obtained.

In Chap. 4 it is discussed how emotions play a role in generating dream episodes from a perspective of internal simulation. Building blocks for this internal simulation are memory elements in the form of sensory representations and their associated emotions. In the presented temporal-causal network model, under influence of associated feeling levels and mutual competition, some sensory representation states pop up in different dream episodes. As a form of emotion regulation the activation levels of both the feelings and the sensory representation states are suppressed by control states. The presented model was evaluated by example simulation experiments.

In Chap. 5 it is discussed how dreaming is used to learn fear extinction. Here fear extinction has been found not to involve weakening of fear associations, as was assumed longer ago, but instead it involves the strengthening of fear suppressing connections that form a counter balance against the still persisting fear associations. So, to regulate fear associations neural mechanisms are used that take care of

strengthening these suppressing connections, as a form of learning of emotion regulation. The presented temporal-causal network model addresses dreaming as internal simulation incorporating memory elements in the form of sensory representations and their associated fear, as in Chap. 4. But this time it is modeled how the regulation of fear that takes place during dream episodes, is strengthened. This adaptation or learning process is modeled as an adaptive temporal-causal network model based on Hebbian learning. The model was evaluated by a number of simulation experiments for different scenarios.

Chapter 6 addresses the role of emotions in rational decision making. Traditionally it has been assumed that emotions can only play a disturbing and non-rational role in decision making. However, more recently it has been found that neurological mechanisms involving emotions play an important role in rational decision making. In this chapter an adaptive temporal-causal network model for decision making based on predictive loops through feeling states is presented, where the feeling states function in a process of valuing of decision options. Hebbian learning is considered for different types of connections in the adaptive model. Moreover, the adaptive temporal-causal network model is analysed from the perspective of rationality. To assess the extent of rationality, measures are introduced reflecting what would be rational for a given environment's characteristics and behaviour. Simulation results and the extents of rationality of different variants of the model over time are discussed. It is shown how during the adaptive process this model for decision making achieves higher levels of rationality.

Part III Yourself and the Others

Part III focuses on persons functioning in a social context. Given that each person has his or her own beliefs, desires, intentions, emotions and still more mental states, it might be expected that social coherence is not often achieved. However, the fact that still often social coherence is observed presents a kind of paradox. This paradox can only be understood by assuming that some neurological mechanisms are responsible for this, and by analyzing more in detail how through such mechanisms influences from the social context affect internal mental processes.

First, in Chap. 7 an overview is presented of a number of recent findings from Social Neuroscience, that form an explanation of how persons can behave in a social manner. For example, shared understanding and collective power are social phenomena that serve as a form of glue between individual persons. They easily emerge and often involve both cognitive and affective aspects. As the behaviour of each person is based on complex internal mental processes involving, for example, own goals, emotions and beliefs, it would be expected that such forms of sharedness and collectiveness are very hard to achieve. Apparently, specific neurological mechanisms are required to tune the individual mental processes to each other in order to enable the emergence of shared mental states and collective behaviour. Having knowledge about these mechanisms provides a basis to modeling corresponding mechanisms in a computational setting. From a neurological perspective, mirror neurons and internal simulation are core concepts to explain the mechanisms underlying such social phenomena. In this chapter it is discussed how based on

such neurological concepts computational mechanisms can be identified to obtain temporal-causal network models for social processes. It is discussed how these models indeed are an adequate basis to simulate the emergence of shared understanding and collective power in groups.

Within a social context the notion of ownership of actions is important. Chapter 8 addresses this notion. It is related to mechanisms underlying self-other distinction, where a self-ownership state is an indication for the self-relatedness of an action and an other-ownership state to an action attributed to someone else. The temporal-causal network model presented in this chapter generates prior and retrospective ownership states for an action based on principles from recent neurological theories. A prior self-ownership state is affected by prediction of the effects of a prepared action as a form of internal simulation, and exerts control by strengthening or suppressing actual execution of the action. A prior other-ownership state plays a role in mirroring and analysis of an observed action performed by another person, without imitating the action. A retrospective self-ownership state depends on whether the sensed consequences of an executed action co-occur with the predicted consequences, and is the basis for acknowledging authorship of actions in social context. It is shown how a number of known phenomena can be obtained as behaviour by the model. For example, scenarios are shown for vetoing a prepared action due to unsatisfactory predicted effects. Moreover, it is shown how poor action effect prediction capabilities can lead to reduced retrospective ownership states, for example, in persons suffering from schizophrenia. This can explain why sometimes the own actions are attributed to others or actions of others are attributed to oneself.

Chapter 9 addresses how in social interaction between two persons usually each person shows understanding of the other person. This may involve both nonverbal and verbal elements, such as bodily expressing a similar emotion and verbally expressing beliefs about the other person. Such social interaction relates to an underlying neural mechanism based on a mirror neuron system. Differences in social responses of individuals can often be related to differences in functioning of certain neurological mechanisms, as can be seen, for example, in persons with a specific type of Autism Spectrum Disorder (ASD). This chapter presents a temporal-causal network model which, depending on personal characteristics, is capable of showing different types of social response patterns based on such mechanisms, adopted from theories on the role of mirror neuron systems, emotion integration, emotion regulation, and empathy in ASD. The personal characteristics may show different variations over time. This chapter also addresses this adaptation over time. To this end it includes an adaptive temporal-causal network model capable of learning social responses, based on insights from Social Neuroscience.

Chapter 10 addresses joint decision making. The notion of joint decision making as considered does not only concern a choice for a common decision option, but also a good feeling about it, and mutually acknowledged empathic understanding about it. In this chapter a temporal-causal network model for joint decision making is presented addressing the role of mutually acknowledged empathic understanding in the decision making. The model is based on principles from recent neurological

theories on mirror neurons, internal simulation, and emotion-related valuing. Emotion-related valuing of decision options and mutual contagion of intentions and emotions between persons are used as a basis for mutual empathic understanding and convergence of decisions and their associated emotions.

In Chap. 11 it is discussed how adaptive temporal-causal network models can be used to model evolving social interactions. This perspective simplifies persons to just one state and expresses the complexity in the structure of the social interactions, modeled by a network. The states can represent, for example, a person's emotion, a belief, an opinion, or a behaviour. Two types of dynamics are addressed: dynamics based on a fixed structure of interactions (modeled by a non-adaptive temporal-causal network model), and dynamics where the social interactions themselves change over time (modeled by an adaptive temporal-causal network model). In the case of an adaptive network model, the network connections change, for example their *weights* may increase or decrease, or connections are added or removed. Both types of dynamics can also occur together. Different types of adaptive social network models are addressed, based on different principles: the *homophily principle* assuming that connections strengthen more when the persons are more similar in their state (the more you are alike, the more you like each other), and the *more becomes more principle* assuming that persons that already have more and stronger connections also attract more and stronger connections. Moreover, it is discussed how dynamics of social interactions can be modeled when (empirical) information over time is available about *actual interaction* between persons (both in the sense of frequency and of intensity), for example, as visible via social media. Based on such information connection weights can be modeled in an adaptive manner: the weights are adapted to the actual interaction.

Part IV Analysis Methods for Temporal-Causal Network Models

Models can be analysed by performing simulation experiments in a systematic manner. For example, it can be found out that under certain conditions a certain state always gets a certain activation level. Moreover, during such experiments values for the parameters of a model can be identified by hand such that for these parameter values the model shows a certain type of behavior. For more complex models such processes may be difficult. In this part some techniques are discussed to achieve this by analysis of the model in different ways.

Chapter 12 addresses the analysis of some types of properties of a temporal-causal network model in an analytical mathematical manner. Properties addressed describe whether some values for the variables exist for which no change occurs (stationary points), whether these variables converge to such a value as a limit value (attracting equilibria), whether variables will show monotonically increasing or decreasing values over time (monotonicity), and whether situations occur in which no convergence takes place but in the end a specific sequence of values is repeated all the time (limit cycle). It is discussed how such analyses can be used for verification of the (implemented) model. Any discrepancies found, suggest there is something wrong in the implementation of the model. In this chapter some methods to analyse such properties of adaptive temporal-causal network models will be

described and illustrated for the Hebbian learning model, and for adaptive connection weights in social network models.

Chapter 13 discusses dynamic properties of processes describing patterns that emerge over time, and how they can be identified and verified in a systematic manner. A process often generates patterns over time that can be described in a temporally more global manner, by expressing temporal relations over longer time periods, in contrast to temporal-causal network model descriptions that specify local mechanisms over small time durations. Such patterns can be considered as emergent phenomena, and it is often a challenge to analyse whether they occur and if so, how their occurrence relates to the local descriptions of underlying mechanisms and their characteristics. Properties describing them have in common that within them references occur to different time points and order relations between time points such as ‘before’ and ‘after’. Moreover, quantifiers over time are used such as expressed by ‘eventually’, ‘always’, ‘during’, ‘for some time point...’, or ‘for all time points...’. Such dynamic properties can be expressed in informal, semiformal and formal ways. Expressing them in a formal numerical-logical format makes it possible to verify whether they hold in some given empirical or simulated scenario in a systematic or even automated manner. This can be helpful in particular if many of such checks have to be done, for example by analysing the effects of a systematic variation of initial values and/or parameters in a simulation experiment.

In Chap. 14 it is discussed how a personalised temporal-causal network model can be obtained that fits well to specific characteristics of persons, and their connections and further context. A model is a close approximation, but always a form of abstraction of a real world phenomenon. Its accuracy and correctness mainly depend on the chosen abstracting assumptions and the values of the parameters in the model. Depending on the complexity of the model, the number of its parameters can vary from just a couple to thousands. These parameters usually represent specific characteristics of the modeled phenomenon, for example, for modeling human processes person-specific characteristics or social interaction characteristics. No values for such parameters are given at forehand. Estimation of parameters for a given model is a nontrivial task. There are many parameter estimation methods available in the literature. In this chapter a number of these methods are briefly discussed.

Part V Philosophical, Societal and Educational Perspectives

In Part V some wider perspectives are addressed. It is discussed how the Network-Oriented Modeling approach relates to historical and philosophical developments concerning dynamics, how it fits in current trends in societal development, and how such a modeling perspective can play a crucial role in an integrative multidisciplinary academic curriculum.

In Chap. 15 it is discussed how dynamics has been a challenging issue in different disciplines since long ago. This issue has been addressed for different domains, in Physics but also in Mathematics, Cognitive Science and Philosophy of Mind. In the development of Physics it has led to notions such as velocity, momentum, kinetic energy and force that drive motion in mechanics. The issue of

dynamics is still out there today, for example, in the domain of Cognitive Science and Philosophy of Mind concerning the physical realism of assumed but not directly physically observable mental states such as desires and intentions that are supposed to drive (physically observable) behaviour. Four cases of dynamics within different traditional disciplines are discussed in this chapter. Similarly, it is shown how in this way causal graphs and transition systems (often used in AI and Computer Science) can be interpreted from a perspective of dynamics. The chapter provides a unified view on the explanation of dynamics across different disciplines. This view is related to the basic assumptions underlying the Network-Oriented Modeling approach based on temporal-causal networks.

Chapter 16 outlines the strong societal development to the integration of more and more smart devices in all aspects of life. Scientific areas addressing this development have names such as Ambient Intelligence, Ubiquitous Computing, Pervasive Computing, Human-Aware Computing or Socially Aware Computing. This development in society often results in integrated complex systems involving humans and technical equipments, also called socio-technical systems. In this chapter it is discussed how in such systems often not only sensor data, but also more and more dynamic computational models based on knowledge from the human-directed sciences such as health sciences, neurosciences, and psychological and social sciences are incorporated. These models enable the environment to perform in-depth analyses of the functioning of observed humans, and to come up with well-informed interventions or actions. It is discussed which ingredients are important to realize this view in a principled manner, among which dynamical models such as temporal-causal network models, and how frameworks can be developed to combine these ingredients to obtain the intended type of systems in practice.

Chapter 17 discusses the design of a curriculum with main focus on human-oriented scientific knowledge and how this can be exploited to develop support for humans by means of advanced smart devices in the daily environment. The aim for this curriculum was to offer a study path for those students with exact talents but with an interest mainly in human processes and society. The curriculum was designed from a problem-oriented perspective in relation to societal problem areas. From human-oriented disciplines scientific knowledge for human processes in such problem areas was obtained. Computational modeling for such human processes plays a central role as an integrating factor in the curriculum. Elements from Ambient Intelligence, Artificial Intelligence, and Informatics are included for design of smart support systems.

Part VI Network-Oriented Modeling: Discussion

Chapter 18 is a discussion in which some of the main issues addressed in the book are briefly reviewed. In particular, the Network-Oriented Modeling approach based on adaptive temporal-causal networks is discussed and how generic and applicable it is as a modeling approach and as a computational paradigm.

References

- J. Aldous, M.A. Straus, Social Networks and Conjugal Roles: a Test of Bott's Hypothesis. *Social Forces* **44**, 576–580, 965–966 (1966)
- S. Aral, L. Muchnik, A. Sundararajan, Distinguishing influence based contagion from Homophily driven diffusion in dynamic networks. *Proc. Natl. Acad. Sci. (USA)* **106**(2), 1544–1549 (2009)
- Aristotle, *Physica* (translated by R.P. Hardie and R.K. Gaye) (350 BC)
- W.R. Ashby, *Design for a Brain*, 2nd edn. (Chapman and Hall, London, 1960). (First edition, 1952)
- W. Barsalou, Simulation, situated conceptualization, and prediction Lawrence. *Phil. Trans. R. Soc. B* **364**, 1281–1289 (2009)
- A. Bechara, H. Damasio, A.R. Damasio, G.P. Lee, Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *J. Neurosci.* **19**, 5473–5481 (1999)
- A. Bechara, H. Damasio, A.R. Damasio, Role of the Amygdala in decision-making. *Ann. N.Y. Acad. Sci.* **985**, 356–369 (2003)
- W. Becker, A.F. Fuchs, Prediction in the oculomotor system: smooth pursuit during transient disappearance of a visual target. *Exp. Brain Res.* **57**, 562–575 (1985)
- R.D. Beer, On the dynamics of small continuous-time recurrent neural networks. *Adapt. Behav.* **3**, 469–509 (1995)
- R.D. Beer, Dynamical approaches to cognitive science. *Trends Cogn. Sci.* **4**, 91–99 (2000)
- A. Bell, Levels and loops: the future of artificial intelligence and neuroscience. *Phil. Trans. R. Soc. Lond. B* **354**, 2013–2020 (1999)
- J. Bickle, *Psychoneural Reduction: The New Wave* (MIT Press, Cambridge, 1998)
- S. Boccaletti, V. Latorab, Y. Morenod, M. Chavez, D.-U. Hwanga, Complex networks: structure and dynamics. *Phys. Rep.* **424**(2006), 175–308 (2006)
- H.W. Bode, *Network Analysis and Feedback Amplifier Design*. Princeton. NJ: Van Nostrand (1945)
- E. Bott, *Family and Social Network: Roles, Norms and External Relationships in Ordinary Urban Families* London: Tavistock Publications (1957)
- D. Byrne, The attraction hypothesis: do similar attitudes affect anything? *J. Pers. Soc. Psychol.* **51**(6), 1167–1170 (1986)
- J.T. Cacioppo, G.G. Berntson, *Social Neuroscience* (Psychology Press, 2005)
- J.T. Cacioppo, P.S. Visser, C.L. Pickett, *Social Neuroscience: People Thinking About Thinking People* (MIT Press, Cambridge, 2006)
- A. Clark, *Being there: Putting Brain, Body, and World Together Again* (MIT Press, 1998)
- F. Crick, C. Koch, Constraints on cortical and thalamic projections: the no-strong-loops hypothesis. *Nature* **391**, 245–250 (1998)
- A.R. Damasio, *Descartes' Error: Emotion, Reason and the Human Brain* (Papermac, London, 1994)
- A.R. Damasio, *The Feeling of What Happens. Body and Emotion in the Making of Consciousness* (Harcourt Brace, New York, 1999)
- A.R. Damasio, *Looking for Spinoza* (Vintage books, London, 2003)
- A.R. Damasio, *Self Comes to Mind: Constructing the Conscious Brain* (Pantheon Books, New York, 2010)
- J. Decety, J.T. Cacioppo (eds.), *Handbook of Social Neuroscience* (Oxford University Press, 2010)
- J. Decety, W. Ickes, *The Social Neuroscience of Empathy* (MIT Press, 2009)
- G. Deliens, M. Gilson, P. Peigneux, Sleep and the processing of emotions. *Exp. Brain Res.* **232**, 1403–1414 (2014). doi:[10.1007/s00221-014-3832-1](https://doi.org/10.1007/s00221-014-3832-1)
- R. Descartes, *The World or Treatise on Light*. Translated version by M.S. Mahoney (1634), <http://www.princeton.edu/~hos/mike/texts/descartes/world/world.htm>
- R.J. Dolan, Emotion, cognition, and behavior. *Science* **298**, 1191–1194 (2002)
- D. Dubois, H. Prade, Possibility theory, probability theory and multiple-valued logics: a clarification. *Ann. Math. Artif. Intell.* **32**, 35–66 (2002)

- D. Dubois, J. Lang, H. Prade, Fuzzy sets in approximate reasoning, Part 2: logical approaches, 1991. *Fuzzy Sets Syst.* **40**, 203–244 (1991) (North-Holland)
- E. Eich, J.F. Kihlstrom, G.H. Bower, J.P. Forgas, P.M. Niedenthal, *Cognition and Emotion* (Oxford University Press, New York, 2000)
- M.S. Elzas, Organizational structures for facilitating process innovation, in *Real Time Control of Large Scale Systems* (Springer, Heidelberg, 1985), pp. 151–163
- P.J. Flory, Network structure and the elastic properties of vulcanized rubber. *Chem. Rev.*, **35**, 51–75 (1944)
- J.P. Forgas, L. Goldenberg, C. Unkelbach, Can bad weather improve your memory? An unobtrusive field study of natural mood effects on real-life memory. *J. Exp. Soc. Psychol.* **45**, 254–257 (2009)
- J.W. Forrester, *World Dynamics*, 2nd edn. (Pegasus Communications, Waltham, 1973), 144 pp
- J.W. Forrester, Lessons from system dynamics modeling. *Syst. Dyn. Rev.* **3**(2), 136–149 (1987)
- N.H. Frijda, A.S.R. Manstead, S. Bem (2000) The influence of emotions on beliefs. in *Emotions and Beliefs: How Feelings Influence Thoughts*, ed. by N.H. Frijda, et al. (Cambridge University Press, 2000), pp. 1–9
- K. Funahashi, Y. Nakamura, Approximation of dynamical systems by continuous time recurrent neural networks. *Neural Networks* **6**, 801–806 (1993)
- V. Gallese, A. Goldman, Mirror neurons and the simulation theory of mind-reading. *Trends Cogn. Sci.* **2**, 493–501 (1998)
- M.S. Gazzaniga (ed.), *The Cognitive Neurosciences*. MIT Press (2009)
- G. Giangiacomo, *Fuzzy Logic: Mathematical Tools for Approximate Reasoning* (Kluwer Academic Publishers, Dordrecht, 2001)
- J. Giles, Computational social science: making the links. *Nature* **488**, 448–450 (2012)
- A.I. Goldman, *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading* (Oxford University Press, New York, 2006), p. 2006
- S. Grossberg, On learning and energy-entropy dependence in recurrent and nonrecurrent signed networks. *J. Stat. Phys.* **1**, 319–350 (1969)
- N. Gujar, S.A. McDonald, M. Nishida, M.P. Walker, A role for REM sleep in recalibrating the sensitivity of the human brain to specific emotions. *Cereb. Cortex* **21**, 115–123 (2011)
- E. Harmon-Jones, P. Winkielman (eds.), *Social Neuroscience: Integrating Biological and Psychological Explanations of Social Behavior* (Guilford, New York, 2007)
- D. Hebb, *The Organisation of Behavior* (Wiley, 1949)
- G. Hesslow, Will neuroscience explain consciousness? *J. Theor. Biol.* **171**(1994), 29–39 (1994)
- G. Hesslow, Conscious thought as simulation of behaviour and perception. *Trends Cogn. Sci.* **6**, 242–247 (2002)
- G. Hesslow, The current status of the simulation theory of cognition. *Brain Res.* **1428**, 71–79 (2012). doi:[10.1016/j.brainres.2011.06.026H](https://doi.org/10.1016/j.brainres.2011.06.026H)
- M. Hirsch, Convergent activation dynamics in continuous-time networks. *Neural Networks* **2**, 331–349 (1989)
- J.J. Hopfield, Neural networks and physical systems with emergent collective computational properties. *Proc. Nat. Acad. Sci. (USA)* **79**, 2554–2558 (1982)
- J.J. Hopfield, Neurons with graded response have collective computational properties like those of two-state neurons. *Proc. Nat. Acad. Sci. (USA)* **81**, 3088–3092 (1984)
- J.C. Hubbard, The Acoustic Resonator Interferometer: I. The Acoustic System and its Equivalent Electric Network. *Phys. Rev.* **38**, 1011 (1931); Erratum *Phys. Rev.* **46**, 525 (1934)
- J.F. Huber, The Arterial Network Supplying the Dorsum of the Foot. *Anatomical Record*, **80**, 373 (1941)
- M. Iacoboni, *Mirroring People: The New Science of How We Connect with Others* (Farrar, Straus & Giroux, New York, 2008)
- P.H. Janak, K.M. Tye, From circuits to behaviour in the amygdala. *Nature* **517**, 284–292 (2015)
- W. James, What is an emotion. *Mind* **9**, 188–205 (1884)
- R.L. Jenison, A. Rangel, H. Oya, H. Kawasaki, M.A. Howard, Value encoding in single neurons in the Human Amygdala during decision making. *J. Neurosci.* **31**, 331–338 (2011)

- J.A.S. Kelso, *Dynamic Patterns: The Self-Organization of Brain and Behavior* (MIT Press, Cambridge, 1995)
- J. Kim, *Philosophy of Mind* (Westview Press, 1996)
- J. Kim, *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation* (MIT Press, Cambridge, 1998)
- B.J. Kuipers, Commonsense reasoning about causality: Deriving behavior from structure. *Artif. Intell.* **24**, 169–203 (1984)
- B.J. Kuipers, J.P. Kassirer, How to discover a knowledge representation for causal reasoning by studying an expert physician, in *Proceedings Eighth International Joint Conference on Artificial Intelligence, IJCAI'83*, Karlsruhe, F.R.G. (William Kaufman, Los Altos, CA, 1983)
- K.S. LaBar, R. Cabeza, Cognitive neuroscience of emotional memory. *Nat. Rev. Neurosci.* **7**, 54–64 (2006)
- G. Lakoff, M. Johnson, *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought* (Basic Books, 1999)
- P.S. Laplace, *Philosophical Essays on Probabilities* (Springer, New York, 1995). Translated by A. I. Dale from the 5th French edition of 1825
- R. Levin, T.A. Nielsen, Disturbed dreaming, posttraumatic stress disorder, and affect distress: a review and neurocognitive model. *Psychol. Bull.* **133**, 482–528 (2007)
- M.D. Lewis, Self-organizing cognitive appraisals. *Cogn. Emotion* **10**, 1–25 (1996)
- G. Loewenstein, J. Lerner, The role of emotion in decision making, in *The handbook of affective science*, ed. by R.J. Davidson, H.H. Goldsmith, K.R. Scherer (Oxford University Press, Oxford, 2003), pp. 619–642
- H.G. Marques, O. Holland, Architectures for functional imagination. *Neurocomputing* **72**, 743–759 (2009)
- W.S. McCulloch, W. Pitts, A Logical Calculus of the Ideas Immanent in Nervous Activity. *Bull. Math. Biophysics* **5**, 115–133 (1943)
- M. McPherson, L. Smith-Lovin, J.M. Cook, Birds of a feather: homophily in social networks. *Annu. Rev. Sociol.* **27**, 415–444 (2001)
- A. Mislove, B. Viswanath, K.P. Gummadi, P. Druschel, You are who you know: inferring user profiles in online social networks, in *Proceedings of WSDM'10*, February 4–6, 2010 (New York City, New York, USA), pp. 251–260
- P.R. Montague, G.S. Berns, Neural economics and the biological substrates of valuation. *Neuron* **36**, 265–284 (2002)
- J.M. Mooij, D. Janzing, B. Schölkopf, From differential equations to structural causal models: the deterministic case, in *Proceedings of the 29th Annual Conference on Uncertainty in Artificial Intelligence (UAI-13)*, ed. by A. Nicholson, P. Smyth. AUAI Press. <http://auai.org/uai2013/prints/papers/24.pdf>. pp. 440–448
- S.E. Morrison, C.D. Salzman, Re-valuing the amygdala. *Curr. Opin. Neurobiol.* **20**, 221–230 (2010)
- M.P. Mundt, L. Mercken, L.I. Zakletskaia, Peer selection and influence effects on adolescent alcohol use: a stochastic actor-based model. *BMC Pediatr.* **12**, 115 (2012)
- E.A. Murray, The amygdala, reward and emotion. *Trends Cogn. Sci.* **11**, 489–497 (2007)
- A. Naudé, D. Le Maitre, T. de Jong, G.F.G. Mans, W. Hugo, Modelling of spatially complex human-ecosystem, rural-urban and rich-poor interactions (2008). https://www.researchgate.net/profile/Tom_De_jong/publication/30511313_Modelling_of_spatially_complex_human-ecosystem_rural-urban_and_rich-poor_interactions/links/02e7e534d3e9a47836000000.pdf
- A. Newell, H.A. Simon, Computer science as empirical inquiry: symbols and search. *Commun. ACM* **19**(3), 113–126 (1976)
- M. Nussbaum (ed.), *Aristotle's De Motu Animalium* (Princeton University Press, Princeton, 1978)
- C. Ouellet, A.A. Benson, The Path of Carbon in Photosynthesis. *Journal of Experimental Botany* **3**, 237–245 (1951)
- O.T. Ousdal, K. Specht, A. Server, O.A. Andreassen, R.J. Dolan, J. Jensen, The human amygdala encodes value and space during decision making. *Neuroimage* **101**, 712–719 (2014)

- E.F. Pace-Schott, A. Germain, M.R. Milad, Effects of sleep on memory for conditioned fear and fear extinction. *Psychol. Bull.* **141**(4), 835–857 (2015)
- J. Pearl, *Causality* (Cambridge University Press, 2000)
- L. Pessoa, On the relationship between emotion and cognition. *Nat. Rev. Neurosci.* **9**, 148–158 (2008)
- L. Pessoa, Emotion and cognition and the amygdala: from “what is it?” to “what’s to be done?”. *Neuropsychologia* **49**, 681–694 (2011)
- G. Pezzulo, M. Candidi, H. Dindo, L. Barca, Action simulation in the human brain: twelve questions. *New Ideas Psychol.* **31**, 270–290 (2013)
- E.A. Phelps, Emotion and cognition: insights from studies of the Human Amygdala. *Annu. Rev. Psychol.* **57**, 27–53 (2006)
- Pineda, J.A. (ed.), *Mirror Neuron Systems: The Role of Mirroring Processes in Social Cognition* (Humana Press Inc., 2009)
- R.F. Port, T. van Gelder, *Mind as Motion: Explorations in the Dynamics of Cognition* (MIT Press, Cambridge, 1995)
- S.M. Potter, What can artificial intelligence get from neuroscience?, in *Artificial Intelligence Festschrift: The Next 50 Years*, ed. by M. Lungarella, J. Bongard, R. Pfeifer (Springer, Berlin, 2007)
- D. Purves, E.M. Brannon, R. Cabeza, S.A. Huettel, K.S. LaBar, M.L. Platt, M.G. Woldorff, *Principles of Cognitive Neuroscience* (Sinauer Associates Inc., Sunderland, 2008)
- A. Rangel, C. Camerer, P.R. Montague, A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience.* **9**, pp. 545–556 (2008)
- G. Rizzolatti, C. Sinigaglia, *Mirrors in the Brain: How Our Minds Share Actions and Emotions* (Oxford University Press, 2008)
- F. Rosenblatt, The Perceptron: A probabilistic Model for Information Storage and Organisation in the Brain. *Psych. Rev.* **65**, 386–408 (1958)
- C.D. Salzman, S. Fusi, Emotion, cognition, and mental state representation in amygdala and prefrontal cortex. *Ann. Rev. Neurosci.* **33**, 173–202 (2010)
- K.R. Scherer, Emotions are emergent processes: they require a dynamic computational architecture. *Phil. Trans. R. Soc. B* **364**, 3459–3474 (2009)
- A. Schurger, S. Uithol, Nowhere and everywhere: the causal origin of voluntary action. *Rev. Phil. Psych.* **6**, 761–778 (2015). doi:[10.1007/s13164-014-0223-2](https://doi.org/10.1007/s13164-014-0223-2)
- C.R. Shalizi, A.C. Thomas, Homophily and contagion are generically confounded in observational social network studies. *Sociol. Methods Res.* **40**(2), 211–239 (2011)
- F. Sotres-Bayon, D.E. Bush, J.E. LeDoux, Emotional perseveration: an update on prefrontal-amygdala interactions in fear extinction. *Learn. Mem.* **11**, 525–535 (2004)
- C.E.G. Steglich, T.A.B. Snijders, M. Pearson, Dynamic networks and behavior: separating selection from influence. *Sociol. Methodol.* **40**, 329–393 (2010)
- J. Storbeck, G.L. Clore, On the interdependence of cognition and emotion. *Cogn. Emot.* **21**, 1212–1237 (2007)
- E. Thelen, L. Smith, *A Dynamic Systems Approach to the Development of Cognition and Action* (MIT Press, Cambridge, 1994)
- L.R.G. Treloar, The Elasticity of a Network of Longchain Molecules. I. *Trans. Faraday Soc.* **39**, 241–246 (1943)
- J. Treur, Temporal factorisation: a unifying principle for dynamics of the world and of mental states. *Cogn. Syst. Res. J.* **8**, 57–74 (2007)
- J. Treur, Dynamic modeling based on a temporal-causal network modelling approach. *Biol. Inspir. Cogn. Archit.* **16**, 131–168 (2016)
- T.W. Valente, *Social Networks and Health: Models, Methods, and Applications* (Oxford University Press, New York, 2010)
- E. van der Helm, J. Yao, S. Dutt, V. Rao, J.M. Saletin, M.P. Walker, REM sleep depotentiates amygdala activity to previous emotional experiences. *Curr. Biol.* **21**(23), 1–4 (2011)
- T. van Gelder, The dynamical hypothesis in cognitive science. *Behav. Brain Sci.* **21**, 615–665 (1998)

- van Gelder and Port, It's about time: an overview of the dynamical approach to cognition, in *Mind as Motion: Explorations in the Dynamics of Cognition*, eds. by R.F. Port, T. van Gelder (MIT Press, Cambridge, 1995), pp. 1–43.
- M.P. Walker, E. van der Helm, Overnight therapy? The role of sleep in emotional brain processing. *Psychol. Bull.* **135**, 731–748 (2009)
- M. Wilson, Six views of embodied cognition. *Psychon. Bull. Rev.* **9**, 625–636 (2002)
- H.V. Westerhoff, A.K. Groen, R.J.A. Wanders, Modern theories of metabolic control and their applications. *Bioscience Reports* **4**, 1–22 (1984)
- N. Wiener, A. Rosenblueth, The mathematical formulation of the problem of conduction of impulses in a network of connected excitable elements, specifically in cardiac muscle, *Arch. Inst. Cardiol. Mexico*. **16**, 202 (1946)
- P. Winkielman, P.M. Niedenthal, L.M. Oberman, Embodied perspective on emotion-cognition interactions, in: *Mirror Neuron Systems: The Role of Mirroring Processes in Social Cognition*, ed. by J.A. Pineda (Humana Press/Springer Science, 2009), pp. 235–257
- S. Wright, Correlation and causation. *J. Agric. Res.* **20**, 557–585 (1921)
- S.S. Yoo, N. Gujar, P. Hu, F.A. Jolesz, M.P. Walker, The human emotional brain without sleep—a prefrontal amygdala disconnect. *Curr. Biol.* **17**, R877–R878 (2007)
- L. Zadeh, Fuzzy sets as the basis for a theory of possibility. *Fuzzy Sets Syst.* **1**, 3–28 (1978). (Reprinted in *Fuzzy Sets and Systems* 100 (Supplement): 9–34, 1999)

Chapter 2

A Temporal-Causal Network Modeling Approach

With Biological, Neurological and Social Processes as Inspiration

Abstract This chapter introduces the Network-Oriented Modeling approach presented in this book. It enables to design complex models in the form of temporal-causal networks, that on the one hand can be described by high level conceptual representations and on the other hand by formal numerical representations of the model. Both representations can easily be translated into each other in a systematic manner. Dedicated software has been developed to support the design of models in a conceptual manner, and to automatically transform them into a numerical and executable format and performing simulation experiments. The modeling format used makes it easy to take into account complex dynamics based on interrelating causal cycles as described by theories and findings about brain processes known from Cognitive, Affective and Social Neuroscience, and complex dynamics based on social interactions. In this chapter also the scope of applicability is discussed in general terms; it is shown that any smooth dynamical system can be modeled by a temporal-causal network. A variety of example models in subsequent chapters illustrates the applicability of the approach in detail.

2.1 Introduction

In this chapter the details of the dynamic Network-Oriented Modeling approach used in this book are discussed. The chapter presents a specific complex systems modeling method which is oriented on networks and is in line with the dynamical systems perspective advocated, for example, by Ashby (1960) and van Gelder and Port (1995). It uses modeling through dynamic networks of temporal-causal relations as a vehicle. This was inspired by the analysis of adaptive causal relations in networks in different physical, biological, neurological, mental and social application domains, among which metabolic networks used as a conceptual structure to model biochemical reactions in cell biochemistry (e.g., Jonker et al. 2002a, b, 2008), networks of mental states used as conceptual structures to model mental processes as considered within Cognitive and Affective Neuroscience and Philosophy of Mind (e.g., Kim 1996), and social networks used as a conceptual

structure to model social interaction; e.g., (Bosse et al. 2015; Sharpanskykh and Treur 2014). The choice for (adaptive) networks of temporal-causal relations provides a modeling approach that can be considered a generic Network-Oriented Modeling approach for dynamics in adaptive complex processes, suitable to obtain a variety of network models: network models for mental processes, biological processes, social processes, and for many other types of processes. The Network-Oriented Modeling approach based on temporal-causal networks has been used in a variety of applications in different domains, and has proven its usefulness (e.g., see the Discussion in Sect. 2.11 of this chapter, and Chaps. 3–11, and the discussion Chap. 18). The current chapter gives a detailed presentation of the modeling approach.

Causal modeling, causal reasoning and causal simulation have a long tradition, in particular, in Artificial Intelligence; e.g., (Kuipers and Kassirer 1983; Kuipers 1984; Pearl 2000). One of the challenges is that causal modeling involving *cyclic paths* in causal graphs poses difficulties; therefore many approaches to causal modeling limit themselves to Directed Acyclic Graphs (DAG's). More in general, to avoid temporal complexity, dynamics is often not addressed in approaches based on causal networks, and the difficulty to allow cyclic paths in a causal network is one consequence of this form of abstraction. Another consequence of abstracting from dynamics is that distinctions in *timing of causal effects* (i.e., how fast causal effects actually are effectuated) cannot be made, whereas sometimes such differences in timing are crucial for the processes modeled by a causal network.

Another nontrivial issue in causal modeling is how to *combine multiple causal impacts* on one state. Apparently, different approaches are possible to address the issue of combining multiple causal impacts (see also Chap. 1, Sect. 1.6). As the applicability of a specific combination rule may depend much on the type of application addressed, and on the type of states within an application, the Network-Oriented Modeling approach based on temporal-causal networks incorporates for each state as a parameter a combination function specifying how multiple causal impacts on this state are aggregated. A number of standard combination functions are available as options and a number of desirable properties of such combination functions have been identified (see Sects. 2.6 and 2.7 below, respectively). These options cover elements from different existing approaches, as mentioned in Chap. 1, Sect. 1.6, varying from approaches considered for reasoning with uncertainty, probability, possibility or vagueness, to approaches based on recurrent neural networks; (e.g., Dubois et al. 1991; Dubois and Prade 2002; Giangiacomo 2001; Zadeh 1978; Hirsch 1989; Hopfield 1982, 1984; Beer 1995). Alternatively, there is still the option to specify any other (non-standard) combination function, preferably taking into account the properties formulated in Sect. 2.7.

The Network-Oriented Modeling approach based on temporal-causal networks described here can be viewed as part of the tradition of causal modeling, but it incorporates the dynamics. This dynamic perspective is based on a continuous time dimension, represented by real numbers. The temporal dimension enables modeling by cyclic causal networks as well, and also timing of causal effects can be modeled in detail. Due to this, causal reasoning and simulation is possible for networks that

inherently contain cycles, such as networks modeling mental or brain states, or networks describing social interaction.

The considered Network-Oriented Modeling approach based on temporal-causal networks has a declarative nature, as, for example, also seen for other approaches developed within Artificial Intelligence, such as knowledge modeling, logical modeling, causal reasoning, model-based diagnosis, or agent modeling. This means that a model description describes (assumed) relations between states (over time) within the domain addressed, and the computational methods for processing or analysis of such relations are separated from the model description itself. Using temporal-causal networks it is relatively easy to design a model at a conceptual, graphical level and to relate the model to scientific literature from a wide variety of disciplines in which such causal relations are also used as a main vehicle to express knowledge. As a particular case, in this way models of mental processes can be related to neuroscientific literature (e.g., from Cognitive Neuroscience, Affective Neuroscience or Social Neuroscience) in which networks are considered that model connections between neurons as a basis for causal relations between the activations of these neurons and the entailed mental states. However, the Network-Oriented Modeling approach based on temporal-causal networks discussed in more detail in this chapter, is much more general and in principle applies to all domains, as in all scientific domains causal relations are commonly used as a way to describe processes.

As discussed above, for one of the types of aggregation of multiple causal impacts, the developed Network-Oriented Modeling approach has similarities to modeling approaches based on neural networks as, for example, described in Hirsch (1989), Hopfield (1982, 1984); Beer (1995). More specifically, the modeling approach adopted here subsumes modeling based on continuous-time recurrent neural networks, an approach advocated by Beer (1995), and inspired, for example, by earlier work in Grossberg (1969), Hopfield (1982, 1984), Funahashi and Nakamura (1993). In Beer (1995) it is claimed that continuous-time recurrent neural networks are an obvious choice for modeling because

- (1) They are the simplest nonlinear, continuous dynamical neural network model
- (2) They are universal dynamics approximators in the sense that, for any finite interval of time, they can approximate the trajectories of any smooth dynamical system on a compact subset of $\mathbb{R}^n$ arbitrarily well (Funahashi and Nakamura 1993)
- (3) They have a plausible neurobiological interpretation

The considered Network-Oriented Modeling approach subsumes the approach advocated by Beer (1995), and the above three advantages also apply to this modeling approach, but in a generalized form:

- (1) Temporal-causal network models are simple dynamical network models
- (2) They are universal dynamics modelers in the sense that, any smooth dynamical system (which by definition is a state-determined system) can be

modeled as a temporal-causal network model (see Sect. 2.9; see also item 6. in the list of desiderata below)

- (3) They have a plausible interpretation in relation to scientific knowledge (from any domain) commonly described by causal relations

The Network-Oriented Modeling approach was developed with a number of more specific desiderata in mind. In this chapter these desiderata are discussed in different (sub)sections. A brief overview of them is:

1. *Modeling dynamics of complex cyclic patterns in real continuous time*

The approach models *dynamics* of simultaneous processes and their often *circular or cyclic patterns* according to *continuous time*, where points and intervals at the time axis are represented by real numbers that correspond to real time points and real time durations.

2. *Models at a conceptual level with relations to physical and biological mechanisms*

The models are described at a *high conceptual* (cognitive, affective, social) *modeling level*, but can be *related* in a transparent manner *to physical and biological mechanisms* underlying the modeled processes, from biologically oriented disciplines, such as (Cognitive, Affective and Social) Neurosciences.

3. *Networks of temporal-causal relations as central element*

By using *networks of temporal-causal relations* as a central modeling element it is facilitated to make use of the large amount of scientific literature in a wide variety of disciplines with knowledge explaining complex processes in terms of *causal relations* between different states.

4. *Design of a conceptual representation as a basis for systematic generation of a detailed numerical representation of a dynamical network model*

Design of a network model can mainly be done at a conceptual level, for example, using *a graphical representation*, and values for a number of parameters and combination functions. On the basis of the conceptual representation of the model a numerical representation can be generated in a systematic manner, or even automatically, and used for simulation experiments and further analysis.

5. *Applicability of multiple computational methods on a given model representation*

The model representations are declarative and have no built in computational methods with them. There is a free choice to apply any computational method on given model descriptions. Such computational methods applied to the model can address different types of tasks such as simulation, or analysis, for example, in the context of verification or validation, or identification of personal, social or other contextual characteristics in the context of empirical information, or analysis or diagnostics of a process, or planning of actions or interventions as part of some application.

6. *Applicability in a wide variety of domains*

The scope of applicability of the Network-Oriented Modeling approach covers mental processes (modeled by temporal-causal networks of mental states), social interaction processes (modeled by temporal-causal social networks), and more. In fact every scientific area in which causal relations are used to describe theories, hypotheses and findings falls within the scope of applicability. This covers practically all scientific domains, as causal explanation is used as a main vehicle almost everywhere in science.

Part of these desiderata have been discussed in Chap. 1, in particular 1.; others will be discussed in more detail in subsequent sections. First, in Sect. 2.2 an introduction of the use of temporal-causal networks is presented. Next, in Sect. 2.3 it will be discussed how findings on neural mechanisms from biologically oriented disciplines such as Cognitive, Affective and Social Neuroscience can be used as a basis and inspiration to model mental and social processes by temporal-causal networks. In Sect. 2.4 the temporal-causal network format used as a basis for dynamical system models is discussed in more detail. Elements used to describe a design of a conceptual representation of a model in the form of such a network are states, weighted causal connections between states, update speed factors, and combination functions to aggregate multiple causal impacts.

In Sect. 2.5 it is shown how on the basis of a declarative conceptual description of a temporal-causal network model, in a systematic manner a detailed declarative numerical representation of this model can be generated which can be used for simulation and further analysis. Such a systematic transformation can be automated by software that has been developed. Section 2.6 discusses a number of standard combination functions that can be used, and in Sect. 2.7 a number of properties of these combination functions are discussed. In Sect. 2.8 it is discussed how to given conceptual or numerical representations different computational methods can be applied, such as network analysis methods, forward or backward simulation methods, model verification, or analysis in the context of empirical data.

In Sect. 2.9 the wide scope of applicability of the Network-Oriented Modeling approach based on temporal-causal networks is discussed. It is shown in some detail how this Network-Oriented Modeling approach enables to model any continuous dynamical system (defined as state-determined system). More specifically, it is shown how any continuous, smooth state-determined system can be modeled according to the Network-Oriented Modeling approach using suitable temporal-causal relations and combination functions to aggregate causal impacts from multiple states on a given state. Section 2.10 shows how adaptive processes can be modeled by an adaptive temporal-causal network model. Finally, Sect. 2.11 is a discussion.

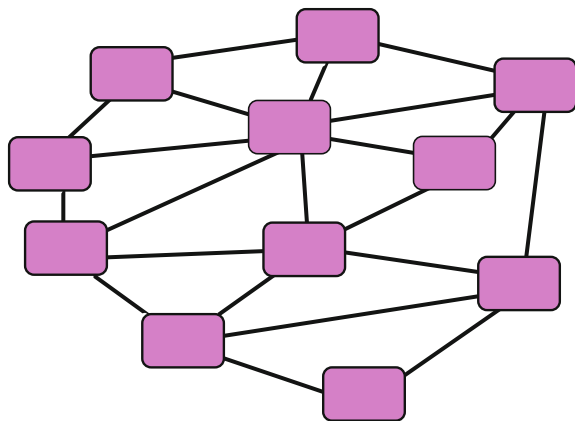
2.2 Modeling Complex Processes by Temporal-Causal Networks

As discussed in Chap. 1, in the domains of Neuroscience, Cognitive Science and Philosophy of Mind the issue of cyclic causal connections between mental states has been recognized. The Network-Oriented Modeling method presented in this book can be used to model the dynamics of cognitive, affective and social interaction processes in an integrative manner by temporal-causal networks. This goes beyond modeling the internal mental processes of individuals; it also covers how processes emerge from individuals interacting with each other. Overall, these processes involve two main aspects:

- a social network structure describing how *persons interact* with each other
- network structures describing specific *individual* mental processes and how they generate their behaviour and social interaction

For the area of social processes inspiration can be obtained from the social network literature within the Social Sciences. A social network model (or short: social network) is a model for social interaction and is often depicted by the connections between persons, as shown in Fig. 2.1, which depicts a graph structure of mutual connections between boxes representing persons. Different types of network models have been distinguished, for example based on fully connected networks, random networks, small world networks and scale-free networks (see also Chap. 11). Many examples of social interaction processes are found in society, that can be described by social networks, and are supported by social media such as Facebook, Whatsapp and Twitter. Within such a social context dynamics are addressed both for mutual influences between states of the different persons involved (*social contagion*), and for the connections between them that may become stronger or weaker over time, or disappear or appear (modeled by *adaptive networks*); e.g., (Sharpanskykh and Treur 2014; see also Chap. 11).

Fig. 2.1 Modeling social interaction processes by a graphical conceptual network representation



The Network-Oriented Modeling approach presented in this book covers such phenomena: a number of network models for these social interaction phenomena will be addressed and analysed, starting in the current chapter in Sect. 2.5.3 and extended in Chaps. 7, 11 and 12. Within a network model for social interaction, social contagion cycles both for cognitive and affective states can play an important role. Such cycles form a basic mechanism that in combination with appropriate internal mental processes enables groups of individuals to develop shared (and empathic) understanding and to make solid joint decisions (see also Chaps. 7 and 10).

For the aspect concerning individual behaviour, internal cognitive and affective processes within each person are addressed. Biological mechanisms discovered in the new and strongly developing disciplines Cognitive, Affective and Social Neuroscience are exploited to obtain scientifically justifiable network models. In this case the internal mechanisms can be described by network models of internal states that causally affect each other (see the network-within-network structure in Fig. 2.2). Note that again network structures are used as a modeling concept, as was the case for social interactions, but in this case they describe interacting mental states.

As discussed in Chap. 1 in more detail, traditionally linear patterns of internal processing are assumed from sensing via sensory processing and preparation to action. However, findings from the neuroscience areas strongly suggest that often parts in the brain are connected by cyclic connections. Such connections are assumed to play an important and regulatory role in many of the brain's processes; e.g., (Bell 1999; Potter 2007). The traditional picture has to be reconsidered, as due

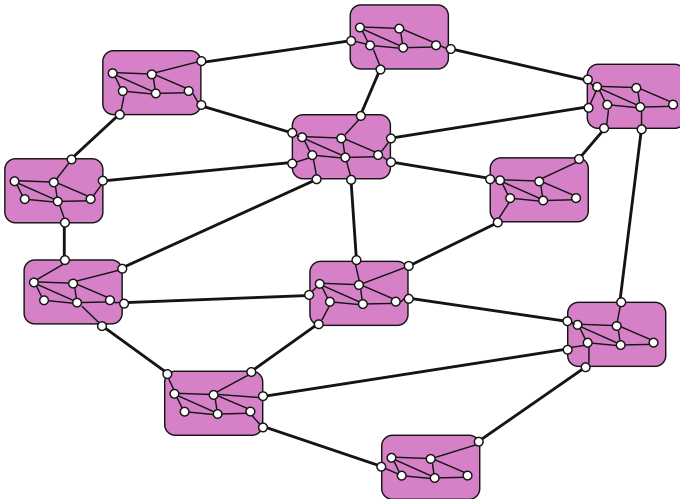


Fig. 2.2 Graphical conceptual network representation of internal processes and social interaction

to mutual, cyclic (e.g., predictive) connections between preparation and sensory states, the processing is often intertwined. Also emotions and feelings are considered to be part of a number of interrelated regulatory cycles, for example, cognitive-affective cycles and emotion regulation cycles. Based on these cycles emotional states emerge over time, and in the meantime affect many other human states and processes.

In the type of network models considered, affective and cognitive processes can be fully integrated in a bidirectional manner: cognitive states lead to associated emotions, and emotions affect cognitive states; e.g., (Scherer 2009). Moreover, for predictive capabilities a core mechanism for *internal simulation* can be used. Internal simulation is internal processing copying an external process, for example another person's mental process. For social functioning this mechanism for internal simulation can be used in combination with another core mechanism: *mirror neuron systems*. Mirror neurons are certain neurons that are activated either for preparation of an action or due to observation of another person performing a corresponding action. These two core mechanisms are the internal basis for the emergence of shared understanding and collective power for a group of persons, based on the social interactions they have. The own individual internal processes are integrated with these mirroring processes (for more details, see Chap. 7).

From Neuroscience it is known that connections within the brain are often adaptive in that their strengths can change over time: *plasticity*. This is a basis for learning; one type of learning that is considered is *Hebbian learning*; e.g., (Hebb 1949; Gerstner and Kistler 2002). This happens when two connected neurons are activated at the same time: such activations lead to strengthening of the connection between them: 'neurons that fire together, wire together'. Such a Hebbian learning principle makes that the states of nodes (within a person) affect the strengths of connections connecting them. This can be applied to learn associations between stimuli and responses, but also, for example, to learn associations of emotions to stimuli or other mental states. Also social interactions are often adaptive. Their structure changes over time, modeled by an *adaptive network*. This can occur in the form of addition and removal of connections, but also in the form of changing the strengths of connections. For example, the homophily principle expresses that persons that are similar in their states will strengthen their connection; see also Chap. 11, or (Sharpanskykh and Treur 2014). Examples of such similarities are having similar beliefs, appreciating the same types of music or movies, or enjoying the same types of activities. Again this makes that the states of nodes (this time within a social context) affect the strengths of the connections connecting them, as in the case of Hebbian learning. Such adaptation principles both for mental states within persons and for interaction between persons within a social context are also covered by the Network-Oriented Modeling approach based on temporal-causal networks presented in this book.

2.3 Exploiting Knowledge About Physical and Biological Mechanisms in Modeling

To design justifiable models with a solid grounding in reality, a clear connection to mechanisms in the physical and biological world is important. Knowledge about such mechanisms is available in various disciplines each addressing certain aspects of reality. However, usually trying to copy such mechanisms directly in a model may easily lead to a model with very high complexity; such a model may lack a desirable extent of transparency. For example, it may become a very complex model covering all neurons, connections and other details of the brain, which is hard to manage. In different scientific disciplines approaches have been developed to handle such complexity by designing higher level models. Sometimes such higher level models still relate well to mechanisms from physical or biological reality although they apply some form of abstraction.

But sometimes higher level models may also relate not so well to such mechanisms, and get a status of being isolated or detached from physical or biological reality. In the latter cases the only relation to reality may be a validation in the form of a comparison of output of the model to data from reality. The internal functioning of such a more detached higher level model may not relate well to mechanisms in reality, and in a sense could be considered as taking place in a black box. The validation provides only a weak justification of such models. For a more solid scientific justification of a higher level model it is important that the internal functioning of the model also relates to scientific knowledge on mechanisms in physical or biological reality. Below this is discussed for two specific scientific domains.

First, in Sect. 2.3.1 the domain of computer science is addressed, where the underlying mechanisms in the physical world concern the functioning of computer hardware. It is discussed how very complex applications concerning processes on a computer can become manageable by describing higher level models of these processes, which still have clear and well-defined relations with the physical mechanisms. Next, in Sect. 2.3.2 the domain of cognitive modeling is addressed. It is discussed how knowledge about neural mechanisms can be used in modeling to obtain models that on the one hand relate in a transparent manner to such mechanisms, but on the other hand are of a higher level of abstraction so that their complexity is manageable.

2.3.1 Addressing Complexity by Higher Level Models Based on Knowledge from Computer Science

The historical development within Computer Science of languages to describe dynamics of processes can be sketched as follows (e.g., Tanenbaum 1976; Knuth 1981; Booch 1991; Kotonya and Sommerville 1998); this has been adopted from

(Jonker et al. 2002). In the early days of Computer Science, languages were developed that describe the dynamics of processes by specifying step by step the (physical) transitions within the physical device called computer. Although for simple processes this may suffice, with a broadening of the scope of applications, these step-by-step descriptions became more and more complex and lacked transparency. Therefore higher level languages that abstract from some of the details of these steps were developed. In the higher level languages a description can be structured in terms of increasingly abstract functional units that cover larger parts of the processes. The result is an increase in the degree of complexity of the phenomena for which transparent descriptions are feasible. This is a way to handle complexity in Computer Science.

Each description in one of the high level languages can be related to (and automatically translated into) lower level descriptions, and ultimately into physical processes that also can be physically executed within a computer. Translation is automated in a generic manner and hidden from the designers who use the higher level languages. In order to perform physical executions in the computer, higher level descriptions are reduced to lower level ones. The benefit for modelers or designers is that by working at the more abstract level of the higher level language, they can keep complexity within the scope of human capabilities, whereas if they would use the lower level descriptions, the task would become too complex and unmanageable.

The Network-Oriented Modeling approach discussed in this chapter is actually also an example of this, as will be shown in subsequent sections. Models can be described in high-level conceptual representations (see Sect. 2.4), which can be automatically transformed into numerical representations (see Sect. 2.5), which in turn can be easily used to obtain computational implementations for simulation in a high level language such as Java, or Matlab, or Excel, which provides an implementation representation which in turn is automatically translated into still lower computer languages until finally physical processes in the computer are performed.

2.3.2 Addressing Complexity by Higher Level Models Based on Knowledge from Neuroscience

Modeling causal relations discussed in neuroscientific literature in a cognitive/affective mental level model does not take specific neurons into consideration but uses more abstract mental states. This abstraction step is a way to make use at the higher modeling level of results from the large and more and more growing amount of neurological knowledge. This method can be considered as lifting neurological knowledge to a higher level of description. In a more detailed manner, in Bickle (1998), pp. 205–208, such a perspective is discussed; for example:

... we can expect that injection of some neurobiological details back into folk psychology would fruitfully enrich the latter, and thus allow development of a more fine-grained folk-psychological account that better matches the detailed functional profiles that neurobiology assigns to its representational states (Bickle 1998, pp. 207–208).

Here Bickle suggests that by relating a (folk) psychological to a neurobiological account, the former can be enriched based on the more detailed description provided by the latter. A similar perspective is advocated by van Gelder and Port (1995). They indicate three main differences between models at the two levels of description: single neuron versus aggregates of neurons, high dimensional versus low dimensional, and short versus long time duration; see:

What is involved in studying processes *at a higher level*? This simple phrase covers a number of different shifts in focus. Most obviously, dynamical cognitive scientists are attempting to describe systems and behaviors that are *aggregates* of vast numbers of systems and behaviors as described at the neural level. Whereas the neuroscientist may be attempting to describe the dynamics of a single neuron, the dynamicist is interested in the dynamics of whole subsystems of the nervous system, comprised of millions, perhaps billions of neurons. Second, the dynamicist obviously does not study this aggregate system by means of a mathematical model with billions of dimensions. Rather, the aim is to provide a *low-dimensional* model that provides a scientifically tractable description of the same qualitative dynamics as is exhibited by the high-dimensional system. Thus, studying systems at a higher level corresponds to studying them in terms of lower-dimensional mathematical models. Third, dynamical cognitive scientists often attempt to describe the neural processes at a larger *time scale* (see Multiple Time Scales, above). The cognitive time scale is typically assumed to lie between roughly a fifth of a second (the duration of an eyeblink) on up to hours and years. It happens to be approximately the range of time scales over which people have awareness of some of their own states and about which they can talk in natural languages. Neuroscientists, by contrast, typically study processes that occur on a scale of fractions of a second (van Gelder and Port 1995, p. 35).

These differences distinguishing lower level neurological models and dynamical models at the higher cognitive/affective level support the view that the latter type of models have advantages over the former in the sense of being easier to handle. At the same time, knowledge about the underlying neural mechanisms often provides crucial input to build realistic and scientifically justifiable models.

2.4 Conceptual Representation of a Temporal-Causal Network Model

To address modeling of the type of cyclic processes as discussed above and in Chap. 1, in this book a dynamical Network-Oriented Modeling perspective is adopted. This modeling perspective covers network models described by a *conceptual representation* of a network of temporal-causal relations in which cycles are allowed, and network models described by a *numerical representation* as often used

in the (numerical) dynamical systems perspective described, for example, in Ashby (1960), Port and van Gelder (1995). These two representations are considered here as two different representations of the same underlying temporal-causal network model. In Sect. 2.5 it is shown how from such conceptual network representation in a systematic manner a corresponding numerical representation for this temporal-causal network model can be obtained.

A description of a temporal-causal network model by a conceptual representation in the first place involves representing in a declarative manner states and connections between them that represent causal relations between states, as assumed to hold for the application domain addressed. What else is needed to describe processes in which such causal relations play their role? In reality not all causal relations are equally strong, so some notion of *strength of a causal connection* is needed. Furthermore, when more than one causal relation affects a state, in which manner do these causal effects combine? Some way to *aggregate multiple causal impacts* on a state is needed. Moreover, not every state has the same extent of flexibility; some states may be able to change fast, and other states may be more rigid and may change more slowly. Therefore, a notion of *speed of change* of a state is used. These three notions are covered by main elements in the Network-Oriented Modeling approach based on temporal-causal networks discussed here:

- *Strength of a causal relation*
Each connection has a *weight value* representing the strength of the causal relation, often between 0 and 1, but sometimes also below 0 (negative effect) or above 1.
- *Combining multiple causal impacts on a state*
For each state (a reference to) a *combination function* is chosen to combine the causal impacts of other states on this state.
- *Speed of change of a state*
For each state a *speed factor* is used to represent how fast a state is changing upon causal impact.

These notions will be discussed in a conceptual manner in Sect. 2.4.1 and illustrated in Sect. 2.4.2.

When any type of function can be used for a combination function, the modeling format is very general and, for example, able to describe any continuous, smooth state-determined system and any set of first-order differential equations, as will be discussed in Sect. 2.9. However, it can also be convenient to make use of a standard type of combination function, so that for a given model it does not need to be specified for each state separately. Specializing the format by committing to a chosen standard type of combination function may make the format less general, but has been shown to still be useful for many purposes. An overview of more often used examples of such standard combination functions can be found in Sect. 2.6.

2.4.1 Conceptual Representations of a Temporal-Causal Network Model

The following standard conceptual representation will be used to describe the temporal-causal network models in this book. The representation has a declarative nature and can be displayed as a graphical representation similar to what is shown in Figs. 2.1, 2.2 and 2.3, or as a matrix representation, as shown, for example, in Table 2.1. A generic description in a graphical representation is shown in Fig. 2.3 and in a matrix representation in Table 2.1; see also Fig. 2.4 for a graphical representation of a specific instance, and Table 2.2 for a matrix representation of this instance. The description or specification of such a conceptual representation of a temporal-causal network model consists of the following elements:

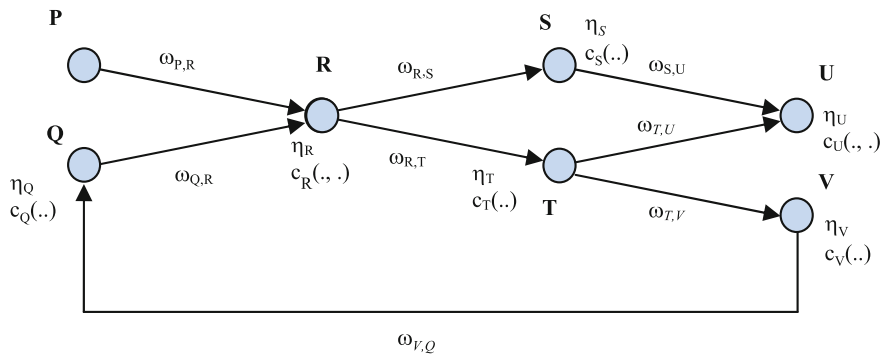


Fig. 2.3 Generic conceptual representation of a temporal-causal network model: with abstract names for states X and Y , connection weights $\omega_{X,Y}$, speed factors η_Y , and combination functions $c_Y(\cdot, \dots, \cdot)$

Table 2.1 Matrix representation of the generic graphical conceptual model representation shown in Fig. 2.3

to from	P	Q	R	S	T	U	V
P			$\omega_{P,R}$				
Q			$\omega_{Q,R}$				
R				$\omega_{R,S}$	$\omega_{R,T}$		
S						$\omega_{S,U}$	
T						$\omega_{T,U}$	$\omega_{T,V}$
U							
V		$\omega_{V,Q}$					
η_Y	η_P	η_Q	η_R	η_S	η_T	η_U	η_V
$c_Y(\dots)$	$c_P(\dots)$	$c_Q(\dots)$	$c_R(\dots)$	$c_S(\dots)$	$c_T(\dots)$	$c_U(\dots)$	$c_V(\dots)$

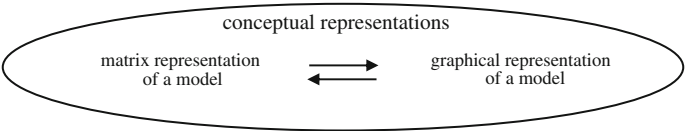


Fig. 2.4 Conceptual representations of a model: matrix representation and graphical representation of a model and systematic mutual transformations into each other

Table 2.2 Matrix representation of the specific model shown in a graphical conceptual representation in Fig. 2.4; this matrix representation can be obtained by instantiation of the generic matrix representation of Table 2.1

to from	P	Q	R	S	T	U	V
P			0.5				
Q			0.3				
R				0.4	0.9		
S						-0.2	
T						0.7	1.0
U							
V		0.8					
η_Y	-	0.1	0.4	0.4	0.4	0.1	0.1
$c_Y(\dots)$	-	id(.)	sum(,..)	id(.)	id(.)	sum(,..)	id(.)

1. A network structure with *names* for *states* Y and *directed connections* between nodes X and Y .
This is the basic conceptual network structure of the model. The nodes indicate states and the connections indicate causal relations between them.
2. For each of the connections from any node X to any node Y a nonzero *weight value* $\omega_{X,Y}$
Such a weight value expresses the strength of the influence of state X on state Y . Also connections from a state to itself are allowed; these can be used to model a form of persistence, for example.
3. For each of the nodes Y a *speed factor* η_Y .
This expresses how fast this state Y will change (update speed).
4. For each of the nodes Y a reference to (which can be given as a name of it) a *combination function* $c_Y(\dots)$.
This function indicates how the multiple impacts from connected states X on state Y are combined into one single impact on state Y . This can be specified, for example, by referring to a sum or scaled sum combination function, a logistic sum combination function, any other standard combination function, or to an own-defined combination function; see Sects. 2.6 and 2.7 for more details of a number of standard combination functions.

In a graphical representation the states and connections are depicted as nodes and arrows (see Figs. 2.3 and 2.4). In a matrix representation the states are at two

axes of a matrix in which each cell can be used to indicate a connection and its weight: the vertical axis indicating the rows is for the ‘from’ states and the horizontal axis indicating the columns is for the ‘to’ states (see Tables 2.1 and 2.2). In a graphical representation the value for the connection weight can be indicated next to the arrow, as shown in Figs. 2.3 and 2.4. However, it is also possible to put a name of the connection or of the connection weight next to the arrow and specify separately what the weight value is for this name. In the matrix representation these elements are written in the cell corresponding to the pair of states involved in the connection. In a graphical representation the speed factors can be written in the pictures close to the states to which they relate. In a matrix representation under the square matrix area for the connections there are two additional rows: one row shows in each column the speed factor η_Y for the state Y indicated for that column, and similarly there is one row for the combination functions $c_Y(\dots)$. More rows can be added for the parameters of such combination functions.

The graph representation depicted in Fig. 2.3 can be transformed into a matrix representation as shown in Table 2.1. Here, first there is a (light blue) square area in which for each pair of states X, Y the connection strength $\omega_{X,Y}$ is indicated in the cell of the row indicated by X and the column indicated by Y .

This matrix represents the same information of a temporal-causal network model as the graphical representation does. Conversely, if a matrix representation of a model is given, it is not difficult to generate an equivalent graphical representation of the same model. The matrix representation and the graphical representation of a temporal-causal network model are equivalent in this sense, and they can be mutually transformed into each other in a systematic manner. In Fig. 2.4 this is summarized. In Sect. 2.5 this picture will be extended by adding a level of numerical representations of the same model and systematic transformations to and from them (see Fig. 2.13).

2.4.2 *More Specific Examples of Conceptual Representations of Temporal-Causal Network Models*

In this section two more specific examples will be discussed to illustrate the type of conceptual representations used. In this context also the practical issue is discussed of how to name the different elements in a model.

An instance of the graphical conceptual representation depicted in Fig. 2.3 is shown in Fig. 2.5. This is an instantiation of the generic structure shown in Fig. 2.3, obtained by specifying values for connection weights and speed factors, and references to specific combination functions. The combination functions chosen here are simple standard combination functions (see also Sect. 2.6): the identity function **id**(.) defined by **id**(V) = V and the sum function **sum**(.,.) defined by **sum**(V_1, V_2) = $V_1 + V_2$. In Table 2.2 the instantiated conceptual matrix representation of the model is shown, as an instance of the generic matrix representation of Table 2.1;

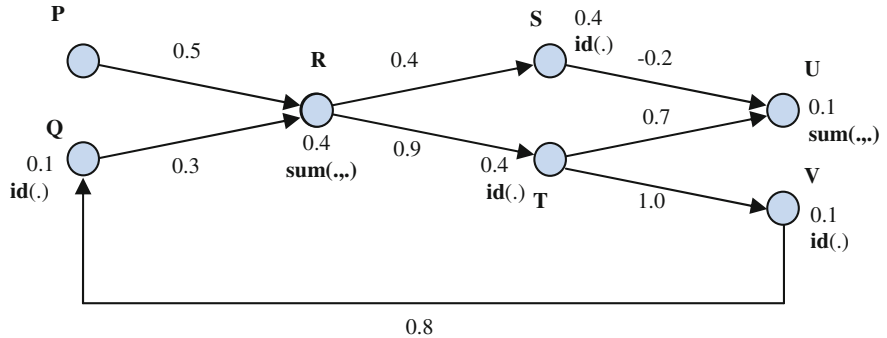


Fig. 2.5 A specific instance of a conceptual representation of a temporal-causal network model: with for states X and Y specific instances for connection weights $\omega_{X,Y}$, speed factors η_Y , and specific references to combination functions $c_Y(\dots)$

note that again the graphical representation and the matrix representation contain the same information.

The states in the above example were not given any meaning. In Fig. 2.6 a more meaningful specific example of a temporal-causal network model is shown. Here some stimulus s is sensed and leads to a sensory representation of s . An example of such a situation is sensing a place to buy icecream (stimulus s); here the sensor state refers to the state of the eyes receiving an image of the place for icecream. The sensory representation in the brain for s formed on the basis of this sensor state in turn affects a preparation state for an action a : the action to go there to get the icecream. This impact leads to a partial activation of this preparation state (with not very high activation level). In turn this preparation state for action a affects the sensory representation of predicted effect e of action a , which is assumed positive: the predicted satisfaction of having the icecream. This state in turn positively affects the preparation state for a , due to which it gets a high activation level: the loop amplifies the activation of the preparation. The high preparation level for action a makes that the execution state for a gets a high activation level, which represents the execution of the action a : actually go to get the icecream.

Note that in Fig. 2.6 the states have no convenient names: they are more like short descriptions and a bit too long to be used as convenient names. To obtain more practical names, first some options for naming are discussed.

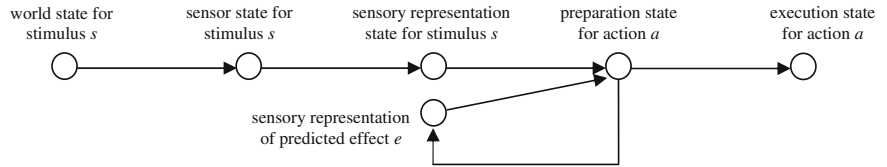


Fig. 2.6 A more specific example of a temporal-causal network model described in a conceptual graphical representation

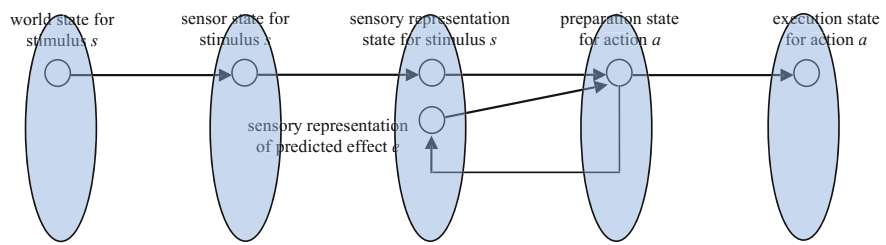


Fig. 2.7 Different groups of states for the example of Fig. 2.6

Naming options: Referring to states

For naming of the states there are different possibilities. How the model works does not depend on the names of the states, so any names can be chosen. However, for more complex models with a larger number of states it is advisable to take some care in giving meaningful names to the states, otherwise the model can easily become difficult to understand for humans. These naming options can be found in the literature:

- Every state a specific state name
For example, P, Q, R, S, T or v, w, x, y, z, or any five self-chosen meaningful words
- Uniform numbered state names
For example, S₁, S₂, S₃, S₄, S₅ or Y₁, Y₂, Y₃, Y₄, Y₅
In some types of literature this type of naming is often chosen. For smaller numbers of states this may work well, but for larger numbers of states this may be not very transparent. For general or theoretical explorations where just any arbitrary model is considered, such a uniform numbered naming format is often used (and actually very useful), where also the number *n* of states is arbitrary; for example, Y₁, ..., Y_{*n*}.
- State names structured according to different groups of states (see Fig. 2.7)
For example, the following five groups of states are often used in the examples.

Group description group	Group name	State names per group
World states	ws	ws _{<i>s</i>}
sensor states	ss	ss _{<i>s</i>}
Sensory representation states	srs	srs _{<i>s</i>} , srs _{<i>e</i>}
Preparation states	ps	ps _{<i>a</i>}
Execution states	es	es _{<i>a</i>}

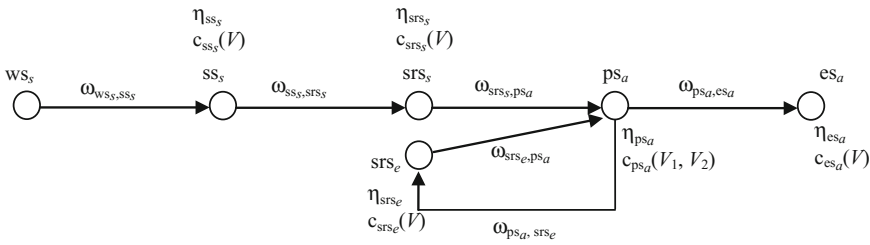
Here for humans sensor states can refer, for example, to states of body parts used to see (eyes), hear (ears), taste (mouth), smell (nose), or feel (skin). In other animals or robots other types of sensor states can be involved as well, for example, for ultrasound, or infrared radiation. This type of naming will be often followed here, as it gives a more transparent conceptual structure to the model. Variants of a notation

Table 2.3 Naming states according to groups

State	Explanation
ws_s	World state for stimulus s
ss_s	Sensor state for stimulus s
srs_X	Sensory representation for X with $X = s$ or $X = e$
ps_a	Preparation state for action a
es_a	Execution state for action a

srs_e are notations such as $srs(e)$ or srs_e . Table 2.3 lists the names of the states based on these groups for the example model of Fig. 2.6. Moreover, in Fig. 2.8 a graphical representation of the example model of Fig. 2.6 is shown based on this naming of states, and in Table 2.4 the same is shown in a matrix representation.

These groups are chosen in such a manner that the model is easily extendable. For example, if more than one stimulus is involved ($s_1, s_2, \dots$) and/or more than one action ($a_1, a_2, \dots$) and/or effects ($e_1, e_2, \dots$), then the groups remain the same and only more instances of states within the groups will be added. Note, that this group structure does not affect how the model actually works, it is only a means to support human understanding of the model. Moreover, the complete description of the

**Fig. 2.8** Referring to states, connection weights, speed factors and combination functions in a standard manner according to groups of states in the graphical conceptual representation of Fig. 2.6**Table 2.4** Matrix representation of the example model shown in Fig. 2.8 in graphical representation

to from	ws_s	ss_s	srs_s	srs_e	ps_a	es_a
ws_s		ω_{ws_s,ss_s}				
ss_s			ω_{ss_s,srs_s}			
srs_s					ω_{srs_s,ps_a}	
srs_e					ω_{srs_e,ps_a}	
ps_a				ω_{ps_a,srs_e}		ω_{ps_a,es_a}
es_a						
η_V	-	η_{ss_s}	η_{srs_s}	η_{srs_e}	η_{ps_a}	η_{es_a}
$c_V(\dots)$	-	$c_{ss_s}(V)$	$c_{srs_s}(V)$	$c_{srs_e}(V)$	$c_{ps_a}(V_1, V_2)$	$c_{es_a}(V)$

Table 2.5 Implicit references to connections and weights

Connection weight	Connection	From	To
ω_{ss_s, srs_s}	(ss_s, srs_s)	Sensor state ss_s for stimulus s	Sensory representation state srs_s of s
ω_{srs_s, ps_a}	(srs_s, ps_a)	Sensory representation state srs_s for stimulus s	Preparation state ps_a for action a
ω_{ps_a, es_a}	(ps_a, es_a)	Preparation state ps_a for action a	Execution state es_a for execution of action a
ω_{ps_a, srs_e}	(ps_a, srs_e)	Preparation state ps_a for action a	Sensory representation state srs_e for e
ω_{srs_e, ps_a}	(srs_e, ps_a)	Sensory representation state srs_e for e	Preparation state ps_a for action a

conceptual matrix representation of the model in Fig. 2.8 is given in Table 2.4. Note that the conceptual graphical representation and the conceptual matrix representation of a model are equivalent: they contain exactly the same information.

Naming options: Referring to connections and their weights

There are two ways to refer to connections: (1) by referring to them implicitly, by only using the states connected by the connection, or (2) by referring to them explicitly by giving the connections names themselves.

1. Referring to connections implicitly by the states they connect

Once the states represented by nodes have their names, there is a simple implicit way to refer to connections. Suppose X and Y are names of states represented by nodes, then the pair notation (X, Y) can be used for the connection from X to Y . Note that this only works when there is at most one connection from one state to another one. In such cases the weight for the connection from X to Y can be denoted by $\omega_{X,Y}$; see Fig. 2.8, Tables 2.4 and 2.5.

However, if state names X and Y are not very simple, this may become complex with many subscripts and sub-subscripts. In cases that states are indicated by numbered variables $Y_1, \dots, Y_n$, the notations ω_{Y_i, Y_j} for connection weights are often simplified to $\omega_{i,j}$.

2. Referring to connections explicitly by giving them their own names

It is also possible to give the connections names themselves. The advantage then is that multiple connections from one state to another can be distinguished by these names. A very simple variant is numbering of the connections: 1, 2, 3, ... (see Fig. 2.9 and Table 2.6). In this case the connection weights can be indicated by $\omega_1, \omega_2, \dots$ third and maybe most informative way of naming a connection is by using a word or small phrase that indicates the process in which the connection is involved. For example, the connection from ps_a to srs_e concerns the process of predicting the effect of the action. Therefore the connection can be named ‘predicting’, and similar for the other connections; see Fig. 2.10 and

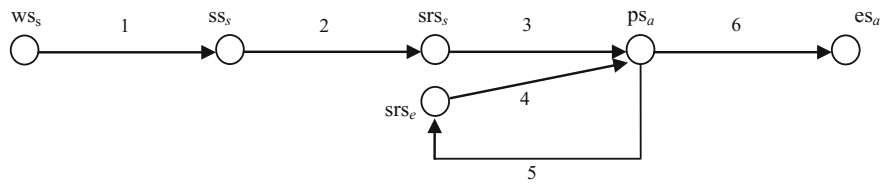


Fig. 2.9 Referring to connections explicitly by assigning numbers to them for the example of Fig. 2.6

Table 2.6 Referring to connections and their weights explicitly by using numbers for the example of Fig. 2.6

Connection weight	Connection name	From state	To state
ω_1	1	ws_s	ss_s
ω_2	2	ss_s	srs_s
ω_3	3	srs_s	ps_a
ω_4	4	srs_e	ps_a
ω_5	5	ps_a	srs_e
ω_6	6	ps_a	es_a

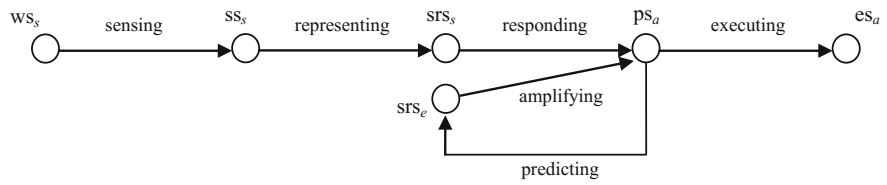


Fig. 2.10 Referring to connections explicitly by assigning names of processes to them, for the example of Fig. 2.6

Tables 2.7 and 2.8. Then weights of connections can be indicated by $\omega_{\text{predicting}}$, and so on. This naming of connections can be made more specific by also including the specific effect e in the name, for example, ‘predicting(e)’ or ‘predicting_ e ’ or ‘predictinge’. Similarly, also the action a may be included in the name of the connection.

Naming options: Referring to speed factors and combination functions

Given the names of nodes representing states, there is a standard way to refer to speed factors and combination functions. Suppose Y is a name of a state represented by a node, then the following notations can be used:

speed factor for state Y η_Y
 combination function for state Y $c_Y(\dots)$

Table 2.7 Referring to connections and their weights explicitly using names indicating processes, for the example of Fig. 2.6

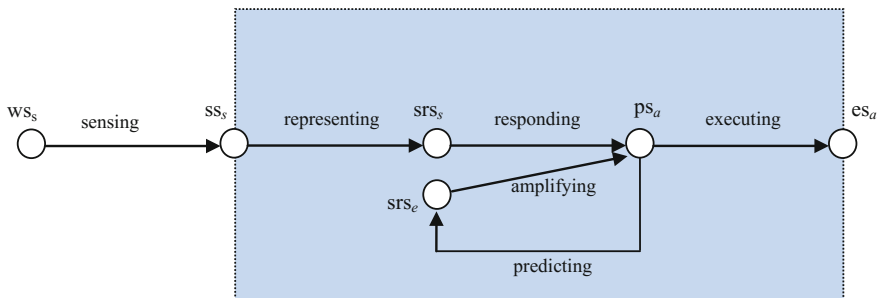
Connection weight	Connection name	From state	To state
ω_{sensing}	Sensing	ws_s	ss_s
$\omega_{\text{representing}}$	Representing	ss_s	srs_s
$\omega_{\text{responding}}$	Responding	srs_s	ps_a
$\omega_{\text{amplifying}}$	Amplifying	srs_e	ps_a
$\omega_{\text{predicting}}$	Predicting	ps_a	srs_e
$\omega_{\text{executing}}$	Executing	ps_a	es_a

Table 2.8 Matrix representation of the example model shown in Fig. 2.10 in graphical representation with meaningful connection names

to from	ws_s	ss_s	srs_s	srs_e	ps_a	es_a
ws_s		ω_{sensing}				
ss_s			$\omega_{\text{representing}}$			
srs_s					$\omega_{\text{responding}}$	
srs_e					$\omega_{\text{amplifying}}$	
ps_a				$\omega_{\text{predicting}}$		$\omega_{\text{executing}}$
es_a						
η_Y	-	η_{ss_s}	η_{srs_s}	η_{srs_e}	η_{ps_a}	η_{es_a}
$c_Y(\dots)$	-	$c_{ss_s}(V)$	$c_{srs_s}(V)$	$c_{srs_e}(V)$	$c_{ps_a}(V_1, V_2)$	$c_{es_a}(V)$

Note that these can be written next to the states in the graph, as shown in Fig. 2.8, but it is often more practical to describe them separately, so that the graph stays more transparent, as for example shown in Figs. 2.9, 2.10 and 2.11.

In Sect. 2.6 a number of standard combination functions and their names are discussed which can be used as more informative names in the model description.

**Fig. 2.11** Indicating boundaries between the internal processes of an individual and processes in the external world

An alternative way of naming for speed factors is by numbering them as $\eta_1, \eta_2, \dots$. In some cases only two different values for speed factors are used, one for fast internal processes and one for slow external processes, which can be indicated by η_{fast} and η_{slow} .

Indicating boundaries of an individual

The boundaries between internal and external processes of an individual person acting in the world (sometimes also called an actor or agent) can be drawn in a graphical representation in the form of a box as shown in Fig. 2.11. Here world states are outside, internal states are inside, and the states for interaction with the world (sensor states, execution states) are on the border of the box. In a matrix representation the rows and columns for internal and external states can be given different colours. Note that the execution state for an action a can lead to an effect of the action in the world (e.g., having the icecream). This can be modeled in the form of an effectuation connection from the execution state es_a to another world state ws_e which in turn can be sensed (this is not depicted in Fig. 2.11). Sensor states have sensing connections from world states to them, and execution states have effectuation connections from them to world states (outside the box). Moreover, sensor states usually have representing connections from them to sensory representation states, and execution states have executing connections from preparation states to them (within the box).

It might seem that indicating a box as in Fig. 2.11 is just a presentation feature by which some colour and emphasis can be added to the picture. But this is not even close to the meaning of such a box. The creation of a boundary between internal processes ('internal milieu', body) and external processes is considered a crucial step in the development of life forms during evolution; for example, see (Bernard 1865; Brewer 1992; Cannon 1932; Damasio 1999, pp. 133–145; Dobbryn and Stuart 2003); see also (Bosse and Treur 2011). The idea is that this boundary can be crossed only by specific processes: from outside to inside by sensor processes (via the sensing states at the boundary), and from inside to outside by execution processes (via the execution states at the boundary). The rest of the boundary is not affectable. As an example, the internal processes for a biological organism are protected against uncontrolled external influences by skin, or bone (protecting the brain), or shell (e.g., like the shell of a sea animal). As another example, a company organised by a 'front office–back office' structure, protects the work going on in the back office against uncontrolled external influences. The front office serves as an interface to the external world, transferring requests for products or services from external to internal and offers for products or services from internal to external.

The fact that the boundary can only be crossed by specific processes via sensor states or execution states can be formulated as a *causal factorisation principle*. This principle expresses that if there is a causal relation between an internal state S_1 and an external state S_2 , then there is an intermediate state R at the boundary with causal

relations to S_1 and S_2 such that the causal relation is between S_1 and S_2 is in fact a composition of these causal relations of S_1 and S_2 with R . In general this factorisation (or interpolation) can be described roughly as a pattern

$$S_1 \rightarrow S_2 \text{ holds} \Rightarrow \text{there exists a state } R \text{ on the boundary such that} \\ S_1 \rightarrow R \text{ and } R \rightarrow S_2 \text{ hold and entail } S_1 \rightarrow S_2$$

Here the two causal relations $S_1 \rightarrow R$ and $R \rightarrow S_2$ are the *factors* of which the causal relation $S_1 \rightarrow S_2$ is composed. The state R can be called a *mediating state* or an *interpolant*. Given the two directions (inward or outward effect) that are possible two more specific causal factorisation principles can be distinguished:

Inward causal factorisation principle: external state S_I causally affecting internal state S_2

If external state S_1 causally affects internal state S_2 , then there exists a sensor state R such that S_1 causally affects R and R causally affects S_2 such that the combined effect of these two causal relations provides the effect of S_1 on S_2	$S_1 \rightarrow S_2$ holds $\Rightarrow$ there exists a sensor state R such that $S_1 \rightarrow R$ and $R \rightarrow S_2$ hold, and $S_1 \rightarrow R$ and $R \rightarrow S_2 \Rightarrow S_1 \rightarrow S_2$
--	---

Outward causal factorisation principle: internal state S_I causally affecting external state S_2

If internal state S_1 causally affects external state S_2 , then there exists an execution state R such that S_1 causally affects R and R causally affects S_2 such that the combined effect of these two causal relations provides the effect of S_1 on S_2	$S_1 \rightarrow S_2$ holds $\Rightarrow$ there exists an execution state R such that $S_1 \rightarrow R$ and $R \rightarrow S_2$ hold, and $S_1 \rightarrow R$ and $R \rightarrow S_2 \Rightarrow S_1 \rightarrow S_2$
--	---

Note that state R can also be a combination of (either sensor or execution) states. In Fig. 2.12 these principles are shown in a graphical form.

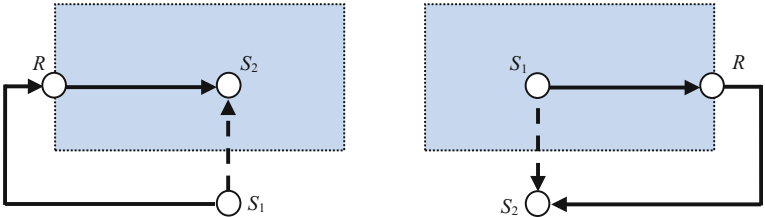


Fig. 2.12 Inward causal factorisation (left hand side) with sensor state R and outward causal factorisation (right hand side) with execution state R

2.5 Numerical Representation of a Temporal-Causal Network Model

In Sect. 2.4 it has been discussed how processes can be modeled in a conceptual, declarative manner in the form of graph or a matrix representation of a network with states as nodes and mutual relations as connections between these nodes. In the current section it is shown how based on such a conceptual representation a declarative numerical representation of the network and its dynamics can be generated. More specifically, it will be shown in detail how from such a conceptual representation in a systematic manner a numerical representation can be obtained, in particular a first-order differential equation or difference equation. In Fig. 2.13 it is shown which types of transformations of model representations are addressed in this chapter. The upper part of this picture shows the two conceptual representations of a model: the graphical representation and the matrix representation. In Sect. 2.4 it has been discussed how they can be transformed into each other, and the upper part of Fig. 2.13 was shown in Fig. 2.4. The lower part of the picture shows two numerical representations. Both types of representations are declarative; they describe (temporal) relations in the domain that is addressed. In the current section it will be shown how a (first-order) numerical difference equation representation can be obtained from a conceptual representation (downward arrow) and from that how in turn a (first-order) numerical differential equation representation can be obtained and vice versa (arrows between numerical representations in the lower part). In Sect. 2.5.1 the systematic transformation from conceptual to numerical representation (downward arrow) will be described in general. In Sect. 2.5.2 it will be illustrated for the example model depicted in Fig. 2.10. Note that in the remainder

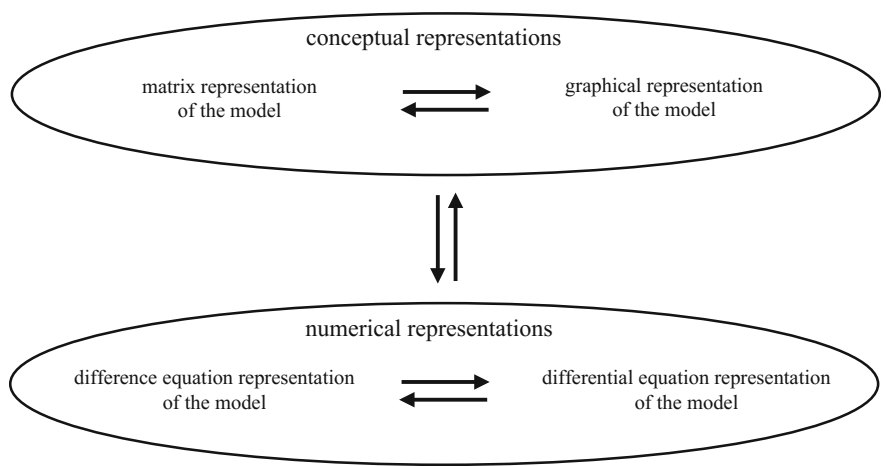


Fig. 2.13 Different representations of a model and transformations between them

of the chapter for a difference equation or differential equation it is always silently assumed first-order.

2.5.1 The Systematic Transformation from Conceptual to Numerical Representation

The transformation from a conceptual representation of a temporal-causal network model to a numerical representation can be used to obtain a basis for simulation and further mathematical analysis. This step is quite systematic and can even be done automatically by a generic software application. Three of such applications have been developed: one in XML and Java, one in Excel, and one as an app on the iPad. During this systematic generation of a numerical representation, the combination functions, that in a conceptual representation are only specified by their name, are incorporated by their mathematical formula. A number of standard combination functions with their formulae are available as a library; see Sect. 2.6. In Table 2.10 an overview is shown of these combination functions and their properties. These formulae can just be taken from this library, but it is also possible to specify a different own-defined, nonstandard combination function. After this numerical representation has been generated in this way, only initial values for the variables are needed in order to perform simulation experiments. Note that parameters are included representing contextual characteristics, such as update speed parameters η_X and connection weights $\omega_{X,Y}$ (see Figs. 2.2 and 2.3), and possibly also parameters in combination functions. To obtain different simulation experiments these also can be varied in order to explore different contexts.

The systematic generation of a numerical representation can be done in the following manner. Here for any state Y and any time point t the (activation) value of Y at time t is denoted by $Y(t)$.

For any of the states Y at each point in time t , each of the values $X_1(t), \dots, X_k(t)$ for the states $X_1, \dots, X_k$ connected toward Y has a causal impact on the value of Y , due to which in principle at the next point in time $t + \Delta t$ the value of Y has changed. For each of the states X_i this impact on Y at time t is proportional both to the value $X_i(t)$ and the connection weight $\omega_{X_i,Y}$ and is defined as

$$\mathbf{impact}_{X_i,Y}(t) = \omega_{X_i,Y} X_i(t)$$

The *aggregated impact* of the multiple impacts $\mathbf{impact}_{X_1,Y}(t), \dots, \mathbf{impact}_{X_k,Y}(t)$ of $X_1(t), \dots, X_k(t)$ on state Y at time t is modeled by a *combination function* $c_Y(\dots)$ (for example, the sum function) as

$$\mathbf{aggimpact}_Y(t) = c_Y(\mathbf{impact}_{X_1,Y}(t), \dots, \mathbf{impact}_{X_k,Y}(t))$$

This aggregated impact is defined on the basis of the network specification in terms of the connection weights $\omega_{X_i,Y}$ and combination functions $c_Y(\dots)$:

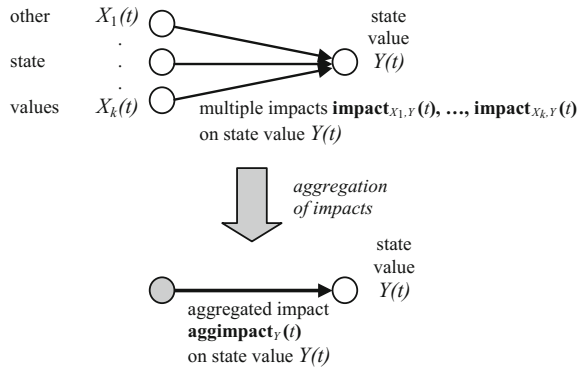
$$\mathbf{aggimpact}_Y(t) = c_Y(\omega_{X_1,Y}X_1, \dots, \omega_{X_k,Y}X_k)$$

Note that here within the combination function an ordering of the arguments (the different impacts by different connections) is used; such an ordering is usually not specified in the conceptual representation of the model. However, many often used combination functions are symmetric, in the sense that the ordering of their arguments does not matter (for example, in a sum, product, max or min function). So, in all of these cases any chosen ordering leads to the same outcome. But in some exceptional cases the order of the arguments may matter; in such cases for the combination function it has to be indicated which argument refers to which connection. Note, however, this is not a temporal order; the multiple impacts are always assumed to work simultaneously, in parallel.

So, the combination function $c_Y(\dots)$ aggregates the multiple impacts $\mathbf{impact}_{X_1,Y}(t), \dots, \mathbf{impact}_{X_k,Y}(t)$ on $Y(t)$ into one single aggregated impact value $\mathbf{aggimpact}_Y(t)$; see Fig. 2.14. Note that also the state Y itself may be included in $X_1, \dots, X_k$, although there are also many cases in which it will not be included. Moreover, as a special case also a combination function can be used for the case of one single impact, i.e., when $k = 1$. Although in such a case it is not literally a process of combination, for convenience also the term combination function is used for a function applied to obtain the (aggregated) impact on Y for this single impact case.

The aggregated impact value $\mathbf{aggimpact}_Y(t)$ at time t has an upward or downward effect on the value of state Y ; it pushes the value of Y up or down, depending on how it compares to the current value of Y . More specifically, this aggregated impact value $\mathbf{aggimpact}_Y(t)$ is compared to the current value $Y(t)$ of Y at t by taking the difference between them (also see Fig. 2.9): $\mathbf{aggimpact}_Y(t) - Y(t)$. If this difference is positive (aggregated impact value $\mathbf{aggimpact}_Y(t)$ at t higher than the current value of Y at t), in the time step from t to $t + \Delta t$ (for some small Δt) the

Fig. 2.14 Aggregation of impacts on a state Y



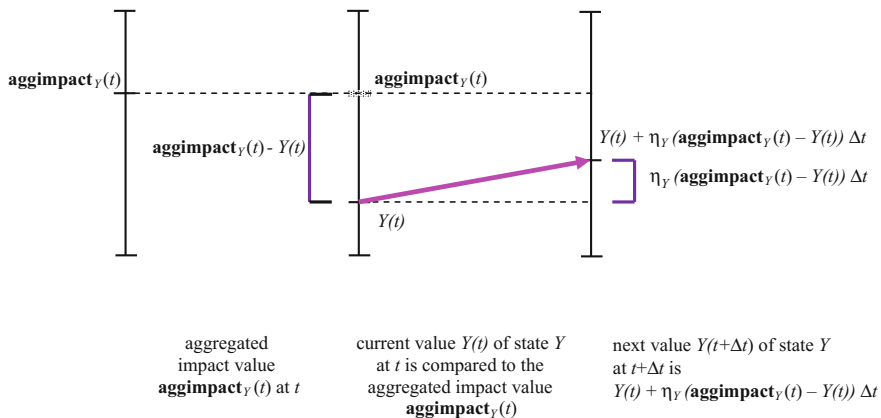


Fig. 2.15 How aggregated impact $\text{aggimpact}_Y(t)$ makes a difference for state $Y(t)$ in the time step from t to $t + \Delta t$

value $Y(t)$ will increase in the direction of the higher value $\text{aggimpact}_Y(t)$. This increase is done proportional to the difference, with proportion factor $\eta_Y \Delta t$: the increase is $\eta_Y [\text{aggimpact}_Y(t) - Y(t)] \Delta t$; see Fig. 2.15.

How fast this increase takes place depends on the speed factor η_Y . For example, when $\eta_Y = 0.9$ and $\Delta t = 0.5$, then a fraction of 0.45 of the difference $\text{aggimpact}_Y(t) - Y(t)$ is added to the value of $Y(t)$. If it holds $\eta_Y = 1$, then the value of Y will adapt to $\text{aggimpact}_Y(t)$ fast (big steps), and if $\eta_Y = 0.1$ it will be much slower (small steps). The same holds for a negative difference $\text{aggimpact}_Y(t) - Y(t)$: in that case the value will decrease in the direction of the lower value $\text{aggimpact}_Y(t)$. The extent to which it is increased depends on the speed factor η_Y .

So the value $Y(t)$ of state Y at t always moves in the direction of the aggregated impact value, and eventually may converge to this value. However, during this convergence process the value of $\text{aggimpact}_Y(t)$ (which itself depends on other states) may change as well, which makes the process still more dynamic.

The numerical process just discussed is summarized by the following *difference equation representation of the dynamical model*:

$$Y(t + \Delta t) = Y(t) + \eta_Y [\text{aggimpact}_Y(t) - Y(t)] \Delta t$$

or, in terms of the combination function $c_Y(\dots)$

$$Y(t + \Delta t) = Y(t) + \eta_Y [c_Y(\omega_{X_1,Y} X_1(t), \dots, \omega_{X_k,Y} X_k(t)) - Y(t)] \Delta t$$

This indicates how for any time point t , the value of Y at a next point in time $t + \Delta t$ can be determined by calculating the right hand side of the equation. This calculation can also be described as

If at time point t
 state X_j has level V_j
 and
 ...
 and state X_k has level V_k
 and state Y has level V
 then after time duration Δt after t state Y will have
 level $V + \eta_Y [c_Y(\omega_{X_1,Y} V_1, \dots, \omega_{X_k,Y} V_k) - V] \Delta t$.

Practically any numerical dynamical modeling language or software environment can be used to model or implement the difference equation format above, for example, Matlab or Excel. Below a summary of the method.

**How to get a numerical representation from a conceptual representation:
 summary of the systematic method**

Given a conceptual representation with

connection weights	$\omega_{X,Y}$
combination functions	$c_Y(\dots)$
speed factors	η_Y

do the following.

For any point in time t and for any state Y

1. Determine the multiple impacts on Y

by multiplying the values $X_1(t), \dots, X_k(t)$ for the states $X_1, \dots$ connected to Y by *connection weights* $\omega_{X_1,Y}, \dots, \omega_{X_k,Y}$:

$$\mathbf{impact}_{X_1,Y}(t), \dots = \omega_{X_1,Y} X_1(t), \dots$$

2. Aggregate the multiple impacts

by applying the *combination function* $c_Y(\dots)$:

$$\mathbf{aggimpact}_Y(t) = c_Y(\mathbf{impact}_{X_1,Y}(t), \dots, \mathbf{impact}_{X_k,Y}(t))$$

3. Incorporate timing

by using the *speed factor* η_Y :

$$\eta_Y [\mathbf{aggimpact}_Y(t) - Y(t)]$$

4. Obtain the difference equation

$$Y(t + \Delta t) = Y(t) + \eta_Y [\mathbf{aggimpact}_Y(t) - Y(t)] \Delta t$$

or fully filled out

$$Y(t + \Delta t) = Y(t) + \eta_Y [c_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) - Y(t)]\Delta t$$

Note that from the above explanation and formulae the following criteria can be obtained for increase, decrease or no change for state X_i :

Increase of Y :

$$\begin{aligned} Y(t + \Delta t) > Y(t) &\Leftrightarrow \mathbf{aggimpact}_Y(t) > Y(t) \\ &\Leftrightarrow c_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) > Y(t) \end{aligned}$$

No change for Y :

$$\begin{aligned} Y(t + \Delta t) = Y(t) &\Leftrightarrow \mathbf{aggimpact}_Y(t) = Y(t) \\ &\Leftrightarrow c_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) = Y(t) \end{aligned}$$

Decrease of Y :

$$\begin{aligned} Y(t + \Delta t) < Y(t) &\Leftrightarrow \mathbf{aggimpact}_Y(t) < Y(t) \\ &\Leftrightarrow c_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) < Y(t) \end{aligned}$$

These criteria will come back in a more general setting in Chap. 12 on mathematical analysis of properties of temporal-causal network models.

The difference equation found above can be rewritten into a different but equivalent numerical differential equation representation as follows. As a first step, by subtracting $Y(t)$ from both sides and dividing by Δt , it can be rewritten into

$$[Y(t + \Delta t) - Y(t)]/\Delta t = \eta_Y [\mathbf{aggimpact}_Y(t) - Y(t)]$$

Note that here the left hand side is a *difference quotient* for $Y(t)$. The limit of this difference quotient for Δt approaching 0 is the differential quotient $\mathbf{d}Y(t)/\mathbf{d}t$ which is the *derivative* $Y'(t)$ of $Y(t)$ as a function of t . So, an alternative (but equivalent) representation is the *differential equation representation of the dynamical model*:

$$\mathbf{d}Y(t)/\mathbf{d}t = \eta_Y [\mathbf{aggimpact}_Y(t) - Y(t)]$$

or

$$\mathbf{d}Y(t)/\mathbf{d}t = \eta_Y [c_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) - Y(t)]$$

Sometimes for the sake of simplicity the variable t is left out of the functions in a differential equation notation, so then it becomes:

$$\mathbf{d}Y/\mathbf{d}t = \eta_Y [\mathbf{aggimpact}_Y - Y]$$

or

$$\mathbf{d}Y/\mathbf{d}t = \eta_Y[\mathbf{c}_Y(\omega_{X_1,Y}X_1, \dots, \omega_{X_k,Y}X_k) - Y]$$

Here the $\mathbf{d}t$ still reveals that for the states in the equation such as $Y, X_1, \dots, X_k$ their activation levels are considered as functions of t . This differential equation format can often be found in the literature. It can easily be (inversely) rewritten into a difference equation format as shown above by replacing $\mathbf{d}Y/\mathbf{d}t$ by $[Y(t + \Delta t) - Y(t)]/\Delta t$, and then multiplying both sides by Δt and adding $Y(t)$.

Note that, like the conceptual representation, also the numerical representation obtained in the form of a differential or difference equation is a declarative description of the addressed domain, this time in the form of numerical (temporal) relations between states in the domain. The way in which these relations are used to perform, for example, computational processes for simulation or analysis is independent of the numerical representation itself. For example, a differential equation can be used to compute the value at a next point in time, from the values at a current time point. However, it is equally well possible to transform the differential equation into a difference equation that can be used to compute the previous value of a state from the values at the current time point. The specific computation applied to the numerical representation can still be chosen; it is independent of the representation itself, it is not coded in this representation.

2.5.2 Illustration of the Transformation for the Example of Fig. 2.10

When the above approach is applied to the example described graphically in Fig. 2.10, with some simple standard combination functions, this results in the following numerical representation of the model. Sometimes the notation $\mathbf{LDP}_X$ or $\mathbf{LP}_X$ is used to indicate the *Local Dynamic Property* describing the dynamics of a state X . They may also be numbered: $\mathbf{LP}_1, \dots, \mathbf{LP}_k$. For a given model each of these local dynamic properties can be specified separately, but together they determine the overall dynamics of the modeled process as a whole. If states and their connections are assumed to be chosen related to some locality criterion, then this illustrates a *locality assumption*: overall dynamics is created based on local mechanisms. However, although in many cases such a locality assumption may hold for the chosen states, the general approach does not commit to such a locality assumption on the states. The approach can be used, for example, equally well to describe dynamics of (communication) actions of operators on Earth that have impact over a large distance on the behaviour of a Mars Explorer.

LP_{ss_s} Sensing a stimulus: determining values for state ss_s

$$ss_s(t + \Delta t) = ss_s(t) + \eta_{ss_s} [c_{ss_s}(\omega_{\text{sensing}} ws_s(t)) - ss_s(t)] \Delta t$$

or

$$dss_s(t)/dt = \eta_{ss_s} [c_{ss_s}(\omega_{\text{sensing}} ws_s(t)) - ss_s(t)]$$

By choosing for the combination function $c_{ss_s}(\cdot)$ the identity combination function **id(.)** defined as **id(V) = V** this becomes

$$ss_s(t + \Delta t) = ss_s(t) + \eta_{ss_s} [\omega_{\text{sensing}} ws_s(t) - ss_s(t)] \Delta t$$

LP_{srs_s} Representing a stimulus: determining values for state srs_s

$$srs_s(t + \Delta t) = srs_s(t) + \eta_{srs_s} [c_{srs_s}(\omega_{\text{representing}} ss_s(t)) - srs_s(t)] \Delta t$$

or

$$dsrs_s(t)/dt = \eta_{srs_s} [c_{srs_s}(\omega_{\text{representing}} ss_s(t)) - srs_s(t)]$$

By again choosing the identity combination function **id(.)** this becomes

$$srs_s(t + \Delta t) = srs_s(t) + \eta_{srs_s} [\omega_{\text{representing}} ss_s(t) - srs_s(t)] \Delta t$$

LP_{ps_a} Preparing for a response: determining values for state ps_a

$$ps_a(t + \Delta t) = ps_a(t) + \eta_{ps_a} [c_{ps_a}(\omega_{\text{responding}} srs_s(t)), \omega_{\text{amplifying}} srs_e(t)) - ps_a(t)] \Delta t$$

or

$$dps_a(t)/dt = \eta_{ps_a} [c_{ps_a}(\omega_{\text{responding}} srs_s(t)), \omega_{\text{amplifying}} srs_e(t)) - ps_a(t)]$$

By choosing for $c_{ps_a}(\cdot, \cdot)$ a sum combination function **sum(.,.)** defined by **sum(V₁, V₂) = V₁ + V₂**, this becomes

$$ps_a(t + \Delta t) = ps_a(t) + \eta_{ps_a} [\omega_{\text{responding}} srs_s(t) + \omega_{\text{amplifying}} srs_e(t) - ps_a(t)] \Delta t$$

LP_{srs_e} Predicting the effect of an action: determining values for state srs_e

$$srs_e(t + \Delta t) = srs_e(t) + \eta_{srs_e} [c_{srs_e}(\omega_{\text{predicting}} ps_a(t)) - srs_e(t)] \Delta t$$

or

$$dsrs_e(t)/dt = \eta_{srs_e} [c_{srs_e}(\omega_{\text{predicting}} ps_a(t)) - srs_e(t)]$$

By choosing the identity combination function **id(.)** this becomes

$$\text{srs}_e(t + \Delta t) = \text{srs}_e(t) + \eta_{\text{srs}_e} [\omega_{\text{predicting}} \text{ps}_a(t) - \text{srs}_e(t)] \Delta t$$

LP es_a Executing an action: determining values for state es_a

$$\text{es}_a(t + \Delta t) = \text{es}_a(t) + \eta_{\text{es}_a} [\text{c}_{\text{es}_a}(\omega_{\text{executing}} \text{ps}_a(t)) - \text{es}_a(t)] \Delta t$$

or

$$\text{des}_a(t)/dt = \eta_{\text{es}_a} [\text{c}_{\text{es}_a}(\omega_{\text{executing}} \text{ps}_a(t)) - \text{es}_a(t)]$$

By choosing the identity combination function **id(.)** this becomes

$$\text{es}_a(t + \Delta t) = \text{es}_a(t) + \eta_{\text{es}_a} [\omega_{\text{executing}} \text{ps}_a(t) - \text{es}_a(t)] \Delta t$$

The numerical relations for the example model obtained above form a declarative, numerical representation of the model. Such a declarative representation is independent of any computational method. In Sect. 2.8 it is shown how a computational simulation method can be applied to the above numerical representation for the example model.

2.5.3 *Illustration of the Modeling Perspective for a Social Contagion Process*

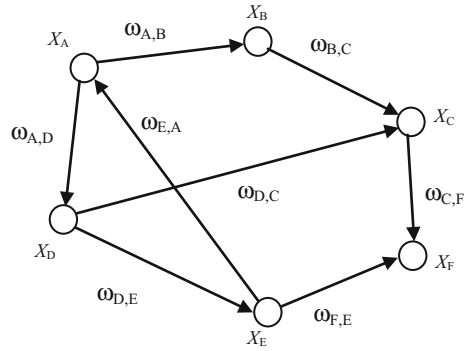
As another illustration, a model is described for contagion effects of multiple individuals on each other due to social interaction. This time each of the persons A is simply modeled by one state X_A , and the interactions between persons are considered by their impact on these states. The state X_A can represent, for example, an emotion or an opinion of person A . First a conceptual representation is shown in Fig. 2.16. The weight of the connection from person A to person B is denoted by $\omega_{A,B}$. Note that more connections may occur than the ones drawn. For example, for every connection from an person X to an person Y there may also exist a connection back from Y to X .

This conceptual representation can be transformed into a numerical representation as shown in Sect. 2.5.1. The following generic numerical representation is obtained, with

$$\text{aggimpact}_B(t) = \text{c}_B(\omega_{A_1,B} X_{A_1}(t), \dots, \omega_{A_k,B} X_{A_k}(t))$$

where the combination function $\text{c}_B(\dots)$ to combine the impacts from the other persons still can be chosen:

Fig. 2.16 Conceptual graphical representation for a simple network view on social contagion



LP1 Social contagion of state X in a social network model

$$\mathbf{d}X_B/\mathbf{d}t = \eta_B [c_B(\omega_{A_1,B}X_{A_1}, \dots, \omega_{A_k,B}X_{A_k}) - X_B]$$

$$X_B(t + \Delta t) = X_B(t) + \eta_B [c_B(\omega_{A_1,B}X_{A_1}(t), \dots, \omega_{A_k,B}X_{A_k}(t)) - X_B(t)]\Delta t$$

Here the symbols are explained as follows:

X_B	person B 's state X
η_B	update speed parameter for state X_B
$c_B(\dots)$	combination function for state X_B
$\omega_{A,B}$	weight of connection from A to B

An often used combination function for aggregation of contagion effects of multiple persons is the *scaled sum function*:

$$\mathbf{ssum}_\lambda(V_1, \dots, V_n) = (V_1 + \dots + V_n)/\lambda$$

where λ is the scaling factor. For this example, where all connection weights $\omega_{A,B}$ are assumed ≥ 0 , this combination function is used for every person B , where the scaling factor $\lambda = \omega_B$ is defined as the sum of the incoming weights for B :

$$\omega_B = \omega_{A_1,B} + \dots + \omega_{A_k,B}$$

So the combination function for any person B is

$$c_B(V_1, \dots, V_n) = \mathbf{ssum}_{\omega_B}(V_1, \dots, V_n) = (V_1 + \dots + V_n)/\omega_B$$

This combination function makes that the aggregated impact from other persons is a weighted average of the individual levels $X_{A_i}(t)$ of state X_{A_i} with weights $\omega_{A_i,B}/\omega_B$ proportional to the connection weights $\omega_{A_i,B}$ and with sum of them 1:

$$\begin{aligned} \mathbf{aggimpact}_B(t) &= c_B(\omega_{A_1,B}X_{A_1}(t), \dots, \omega_{A_k,B}X_{A_k}(t)) \\ &= (\omega_{A_1,B}X_{A_1}(t) + \dots + \omega_{A_k,B}X_{A_k}(t))/\omega_B \\ &= (\omega_{A_1,B}/\omega_B)X_{A_1}(t) + \dots + (\omega_{A_k,B}/\omega_B)X_{A_k}(t) \end{aligned}$$

with

$$\omega_{A_1,B}/\omega_B + \dots + \omega_{A_k,B}/\omega_B = 1$$

The interpretation of this combination function is that a person B adapts the level of his or her state X_B to what is observed by B as the average level of state X_A for all persons A that have interaction to B ; this can result in adapting to the majority in the population described by the network. However, if the connections only relate to a small and relatively isolated part of the population, the average over this part can deviate from the average of the population as a whole.

If this scaled sum combination function is chosen, the numerical representation of the contagion model becomes:

LP1 Social contagion of state X in a social network model with scaled sum combination function

$$\begin{aligned} dX_B/dt &= \eta_B[(\omega_{A_1,B}X_{A_1} + \dots + \omega_{A_k,B}X_{A_k})/\omega_B - X_B] \\ X_B(t + \Delta t) &= X_B(t) + \eta_B[(\omega_{A_1,B}X_{A_1}(t) + \dots + \omega_{A_k,B}X_{A_k}(t))/\omega_B - X_B(t)]\Delta t \end{aligned}$$

Here the symbols are explained as follows:

- X_B person B 's state X
- η_B update speed parameter for state X_B
- $c_B(\dots)$ combination function for state X_B
- $\omega_{A,B}$ weight of connection from A to B
- ω_B sum of incoming weights = $\omega_{A_1,B} + \dots + \omega_{A_k,B}$

Figure 2.17 shows an example simulation for this social contagion model. In Chap. 12 a mathematical analysis will be discussed, for example it is shown that all

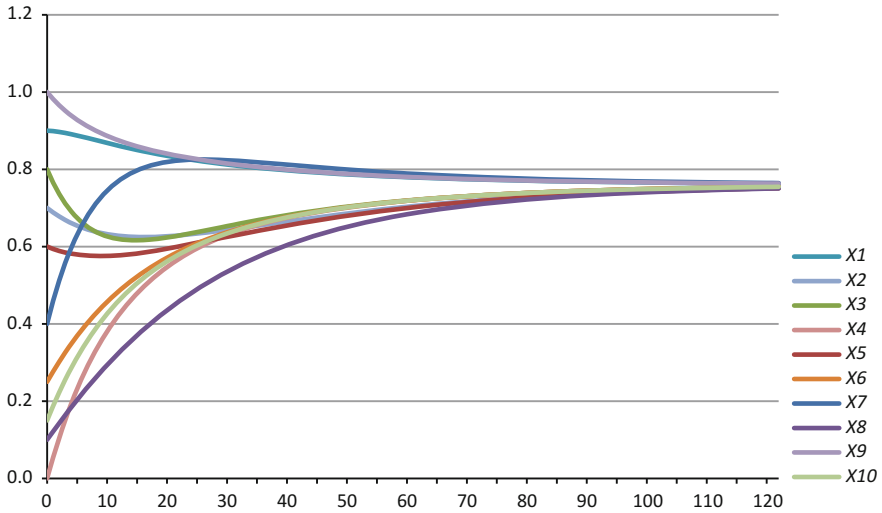


Fig. 2.17 Example simulation of social contagion for 10 persons

state values will always converge to one and the same value, as long as the network is strongly connected. In Chaps. 11 and 12 also adaptive extensions of this model will be discussed.

2.6 Standard Combination Functions

In this section a number of often used examples of standard combination functions are discussed. For an overview, see Table 2.10. These standard combination functions can be used during the design of a model as immediately available building blocks. In software environments that automatically generate a numerical representation out of a model in conceptual graphical or conceptual matrix representation, these standard combination functions can be used without implementing them first, as they are already offered by the software. In the design of a model only a reference to the name of the chosen standard combination function is needed, and sometimes values for some parameters of this combination function.

2.6.1 Basic Standard Combination Functions

Standard combination functions can be of two types. For the *basic* type they are directly based on simple functions such as **sum**, **product**, **max**, **min** or a simple **logistic** function. These are discussed in this Sect. 2.6.1. However, also standard combination functions of a more *advanced* type can be considered. In this type of functions combinations of simple functions are used. These will be discussed in Sect. 2.6.2. In Sect. 2.7 a number of properties that are relevant for combination functions will be discussed.

The basic standard combination functions considered are SCF1 to SCF4 described below.

SCF1 Sum and identity combination function

In the sum combination function, the different single impacts are simply added to obtain the aggregate impact:

$$c(V_1, \dots, V_k) = \mathbf{sum}(V_1, \dots, V_k) = V_1 + \dots + V_k$$

Note that when $k = 1$, this is just the identity function denoted by **id(.)** and defined as **id**(V) = V . Moreover, note that in cases that this combination function is used for a state X_i with $X_1, \dots, X_k$ connected toward X_i with $\sum_j \omega_{j,i} = 1$, then the sum combination function makes a weighted average of the values of states $X_1, \dots, X_k$:

$$\mathbf{sum}(\omega_{1,i}X_1, \dots, \omega_{k,i}X_k) = \omega_{1,i}X_1 + \dots + \omega_{k,i}X_k$$

Note that in the example shown in Figs. 2.10 and 2.11, for the preparation state indeed the sum of the incoming connections was 1 (both were 0.5), so in that case indeed a weighted average was used.

SCF2 Product-based combination functions

The product combination function is defined as a multiplication

$$c(V_1, \dots, V_k) = \mathbf{product}(V_1, \dots, V_k) = V_1 * \dots * V_k$$

The complementary product combination function (for values between 0 and 1) is defined as

$$c(V_1, \dots, V_k) = \mathbf{cproduct}(V_1, \dots, V_k) = 1 - (1 - V_1) * \dots * (1 - V_k)$$

The use of these product combination functions may get inspiration from a probability view for cases in which different impacts are fully independent, in which case the probability of the combination is the product of the probabilities.

Note that for $V_i \leq 1$, the first variant provides a result lower than (or at most as low as) each of the V_i ; the second, complementary variant provides a result higher than (or at least as high as) each V_i :

$$\begin{aligned} \mathbf{product}(V_1, \dots, V_k) &\leq \mathbf{min}(V_1, \dots, V_k) \\ \mathbf{cproduct}(V_1, \dots, V_k) &\geq \mathbf{max}(V_1, \dots, V_k) \end{aligned}$$

Note that in cases that this combination function is used for a state X_i with $X_1, \dots, X_k$ connected toward X_i then the product combination function provides a product of the values of $X_1, \dots, X_k$ with a product of the values of $\omega_{j,i}$:

$$\mathbf{product}(\omega_{1,i}X_1, \dots, \omega_{k,i}X_k) = \omega_{1,i}X_1 * \dots * \omega_{k,i}X_k = (\omega_{1,i} * \dots * \omega_{k,i}) * (X_1 * \dots * X_k)$$

The complementary product combination function:

$$\mathbf{cproduct}(\omega_{1,i}X_1, \dots, \omega_{k,i}X_k) = 1 - (1 - \omega_{1,i}X_1) * \dots * (1 - \omega_{k,i}X_k)$$

SCF3 Minimum or maximum combination function

The **min** and **max**-based combination function are defined as follows.

$$c(V_1, \dots, V_k) = \mathbf{min}(V_1, \dots, V_k) \quad \text{the minimal value among the } V_i$$

$$c(V_1, \dots, V_k) = \mathbf{max}(V_1, \dots, V_k) \quad \text{the maximal value among the } V_i$$

The use of these combination functions may get inspiration from approaches from fuzzy logic and possibility theory (Zadeh 1978), where also such operators are used to combine different sources of impact.

Note that the first minimum variant provides a result lower than (or at most as low as) each of the X_i ; the second, maximum variant provides a result higher than (or at least as high as) each X_i .

The minimum combination function has an effect similar to an *and* operator: it only gives a high aggregated impact if *all single impacts* are high (X_1 and ... and X_k have a high level). Similarly the maximum combination function has an effect similar to an *or* operator: it already gives a high aggregated impact if *at least one of the single impacts* is high (X_1 or ... or X_k has a high level).

In cases that this combination function is used for a state X_i with $X_1, \dots, X_k$ connected toward X_i then these functions work as follows:

$$c(\omega_{1,i}X_1, \dots, \omega_{k,i}X_k) = \mathbf{min}(\omega_{1,i}X_1, \dots, \omega_{k,i}X_k)$$

$$c(\omega_{1,i}X_1, \dots, \omega_{k,i}X_k) = \mathbf{max}(\omega_{1,i}X_1, \dots, \omega_{k,i}X_k)$$

SCF4 Simple logistic sum combination function

The logistic sum combination function has two closely related variants, the simple variant SCF4 and the more advanced variant SCF5 (see below). In these functions τ is a threshold parameter and σ a steepness parameter. The simple logistic function SCF4 is defined as:

$$c(V_1, \dots, V_k) = \mathbf{logistic}(V_1, \dots, V_k) = 1/(1 + e^{-\sigma(V_1 + \dots + V_k - \tau)})$$

To indicate the dependence of σ and τ sometimes these are used as subscripts: $\mathbf{logistic}_{\sigma,\tau}(V_1, \dots, V_k)$.

In cases that this combination function is used for a state X_i with $X_1, \dots, X_k$ connected to X_i then this function works as follows on the X_i :

$$\mathbf{logistic}(\omega_{1,i}X_1, \dots, \omega_{k,i}X_k) = 1/(1 + e^{-\sigma(\omega_{1,i}X_1 + \dots + \omega_{k,i}X_k - \tau)})$$

In Fig. 2.18 the simple logistic combination function $\mathbf{logistic}(V_1, \dots, V_k)$ is depicted with $W = V_1 + \dots + V_k$ on the horizontal axis, for a number of values of the steepness parameter.

Note that it holds $c(0, \dots, 0) = 1/(1 + e^{\sigma\tau})$, and this is nonzero, which may be considered an artifact of the model which lacks plausibility: it creates activation from no activation. In Fig. 2.18, this is visible especially in the highest graph for $\sigma = 5$, but also in the other cases it is (although closer to) still not zero. This less desirable phenomenon may suggest to only use the simple logistic sum combination function for high steepness values, for example, 20 or higher (this also may depend on the threshold τ). However, then it gets a more step function like character and lacks the more gradual increasing behavior as shown for steepness 10 and lower. If such a gradual increase is aimed for, the advanced logistic sum function discussed below may be a better choice. Moreover, note that for very high steepness value σ this logistic combination function provides a continuous approximation of a binary threshold-based function that gives 0 for $V_1 + \dots + V_k$ below the threshold τ and 1 for $V_1 + \dots + V_k$ above this threshold (e.g., see Fig. 2.18 lowest graph).

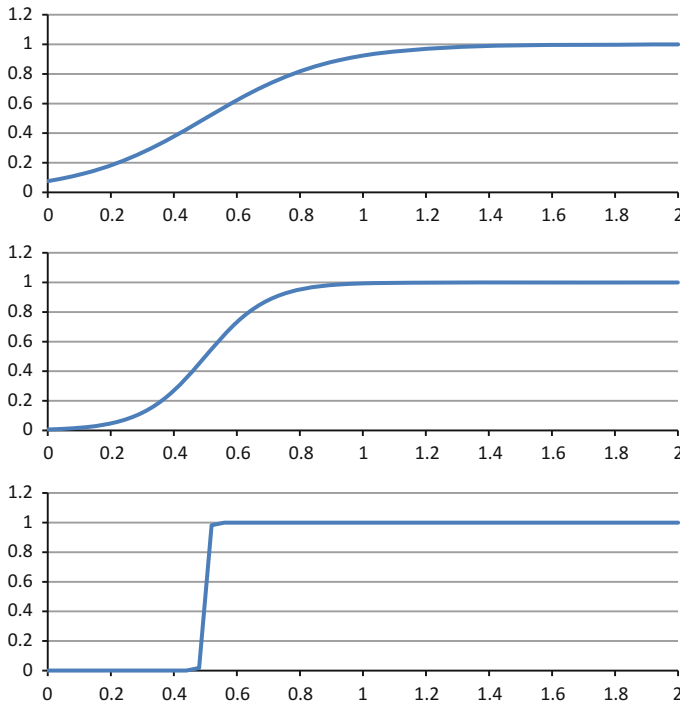


Fig. 2.18 Simple logistic function for threshold $\tau = 0.5$ and four different values of steepness: $\sigma = 5$, $\sigma = 10$, $\sigma = 200$

2.6.2 Building More Complex Standard Combination Functions

From relatively simple or basic combination functions as described in Sect. 2.6.1 more complex combination functions can be built by considering some combinations, modifications or transformations on them. This can be done, for example, by adding or multiplying constants to one of them, or by forming linear combinations (such as a weighted sum) of some of them.

SCF5 Advanced logistic sum combination function

Recall that in the simple logistic variant SCF4 it holds $\mathbf{slogistic}(0, \dots, 0) = 1/(1 + e^{\sigma\tau})$, and this is nonzero, which is an undesirable property as it creates in an unintended manner activation out of no activation (see also property CFP1 in Sect. 2.7). This issue is compensated for in the advanced variant. This advanced variant is obtained by first subtracting the value at 0 from it:

$$1/(1 + e^{-\sigma(V_1 + \dots + V_k - \tau)}) - (1/(1 + e^{\sigma\tau}))$$

Then a function is obtained that for $V_1 + \dots + V_n = 0$ gives value 0, so that problem is solved (this new combination functions satisfies property CFP1 of Sect. 2.7). However, now the maximal value can be at most

$$1 - (1/(1 + e^{\sigma\tau})) = (1 + e^{\sigma\tau})/(1 + e^{\sigma\tau}) - (1/(1 + e^{\sigma\tau})) = \\ ((1 + e^{\sigma\tau}) - 1)/(1 + e^{\sigma\tau}) = e^{\sigma\tau}/(1 + e^{\sigma\tau}) = 1/(1 + e^{-\sigma\tau})$$

which is strictly lower than 1. But this can be compensated for by dividing the whole expression by this value. This provides a function that gives values ranging from 0 to 1:

$$[1/(1 + e^{-\sigma(V_1 + \dots + V_k - \tau)}) - (1/(1 + e^{\sigma\tau}))]/[1/(1 + e^{-\sigma\tau})] = \\ [(1/(1 + e^{-\sigma(V_1 + \dots + V_k - \tau)})) - (1/(1 + e^{\sigma\tau}))](1 + e^{-\sigma\tau})$$

So, the more **advanced logistic sum combination function** obtained is defined as:

$$c(V_1, \dots, V_k) = \mathbf{alogistic}(V_1, \dots, V_k) \\ = [(1/(1 + e^{-\sigma(V_1 + \dots + V_k - \tau)})) - (1/(1 + e^{\sigma\tau}))](1 + e^{-\sigma\tau})$$

To indicate the dependence of σ and τ sometimes these are used as subscripts:

$$\mathbf{alogistic}_{\sigma, \tau}(V_1, \dots, V_k)$$

For this advanced logistic function graphs are shown in Fig. 2.19. Note that only for the upper graph for $\sigma = 5$ a difference is visible with the corresponding upper graph in Fig. 2.18 for $\sigma = 5$. The other two graphs in Fig. 2.19 for $\sigma = 10$ and $\sigma = 200$ also differ from the corresponding graphs in Fig. 2.18, but these differences are so small that they are (almost) not visible. This indicates that the simple variant can be used as a suitable approximation of the advanced one when $\sigma\tau$ is large enough, e.g., $\sigma\tau \geq 20$. However, using high steepness values the model shows more abrupt behavior, for example, when one of the impact values increases from below to above the threshold value. In some cases such behaviour may be realistic and intended, but in other cases it is not. To model more smoothly behaving processes, often lower values of the steepness σ are needed, in which case the advanced logistic function is more appropriate.

SCF6 Scaled sum combination function

In some cases it is useful to apply a scaling factor to the sum combination function by dividing it by some *scaling factor* λ :

$$c(V_1, \dots, V_k) = \mathbf{ssum}_{\lambda}(V_1, \dots, V_k) = (V_1 + \dots + V_k)/\lambda$$

This can be used, for example, to guarantee that values are kept between 0 and 1 for cases that $\omega_{j,i} \geq 0$ for all j and i (property CFP5 in Sect. 2.7). If $\omega_{j,i} \geq 0$ for all

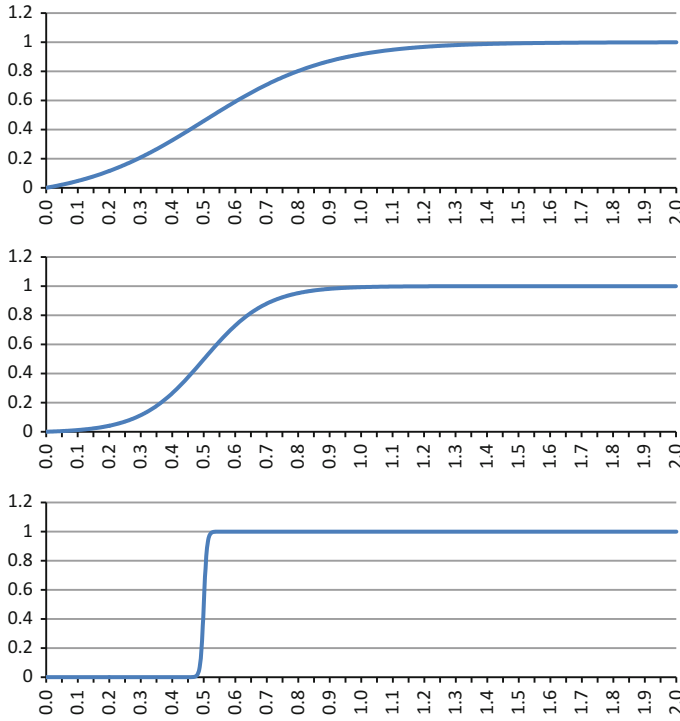


Fig. 2.19 Advanced logistic function for threshold $\tau = 0.5$ and three different values of steepness: $\sigma = 5$, $\sigma = 10$, $\sigma = 200$

j and i and $X_j \leq 1$ for all j , the sum of $\omega_{1,i} X_1 + \dots + \omega_{k,i} X_k$ can reach the value $\omega_{1,i} + \dots + \omega_{k,i}$, which can be more than 1, depending on the values of these connection weights $\omega_{j,i}$. By scaling (normalizing) a sum function, a function is obtained that keeps values between 0 and 1 for cases that $\omega_{j,i} \geq 0$ for all j ; if as a scaling factor

$$\lambda_i = \omega_{1,i} + \dots + \omega_{k,i}$$

is taken, then this combination function indeed keeps values between 0 and 1. In fact, this scaled sum combination function provides as outcome a *weighted average* of the values of X_1 to X_k : for this scale factor λ_i it holds

$$\begin{aligned} \text{ssum}_{\lambda,i}(V_1, \dots, V_k) &= (\omega_{1,i}X_1 + \dots + \omega_{k,i}X_k) / (\omega_{1,i} + \dots + \omega_{k,i}) \\ &= (\omega_{1,i} / (\omega_{1,i} + \dots + \omega_{k,i}))X_1 + \dots + (\omega_{k,i} / (\omega_{1,i} + \dots + \omega_{k,i}))X_k \end{aligned}$$

which is a weighted sum where the sum of the weights is 1:

$$\omega_{1,i}/(\omega_{1,i} + \dots + \omega_{k,i}) + \dots + \omega_{k,i}/(\omega_{1,i} + \dots + \omega_{k,i}) = 1$$

Note that when one or more of the $\omega_{j,i}$ is negative, the scaled sum combination function may result in negative values, and using the sum $\omega_{1,i} + \dots + \omega_{k,i}$ as scaling factor, this scaling factor can be close or even equal to 0 and therefore it may easily produce results that are above 1 or even infinite. Due to this, for cases with negative connection weights the scaled sum combination function with the sum of the weights as scaling factor may not be a recommendable choice of combination function. Instead, a different scaling factor might be considered, for example, the sum of the absolute values of the weights or the sum of only the positive weights, but these options may also not work so well when multiple or most weights are negative. The first option, for example, will always result in aggregated values that are much lower than 1, and the second option may result in negative values far beyond -1 .

For certain settings of the parameters the (scaled) sum function and the advanced logistic sum function approximate each other. For example, in the upper graph in Fig. 2.20 this is shown for the sum function ($\lambda = 1$) and the advanced logistic sum function for $\sigma = 2.402$ and $\tau = 0.142$; in the lower graph the same is shown for $\lambda = 2$ and $\sigma = 1.084$ and $\tau = 0.143$. Note that for this approximation rather low steepness values are used, what would not be possible using the simple logistic function due to its problem with zero activation. An overview for some more values is shown in Table 2.9. Here α is the slope of the linear function, λ is the

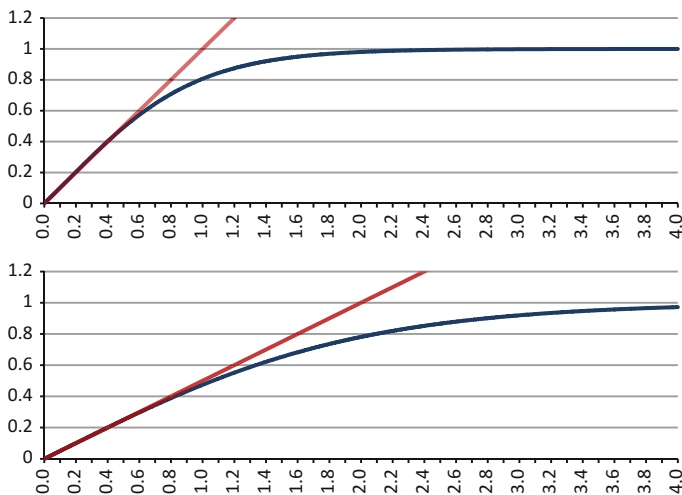


Fig. 2.20 Approximation between scaled sum and advanced logistic sum function: upper graph $\lambda = 1$, $\sigma = 2.402$ and $\tau = 0.142$, lower graph $\lambda = 2$, $\sigma = 1.084$ and $\tau = 0.143$

Table 2.9 Corresponding settings for approximation between scaled sum and advanced logistic sum

λ	4.000	3.000	2.000	1.000	0.667
α	0.25	0.33333	0.500	1.000	1.500
σ	0.5192	0.7009	1.084	2.402	4.137
τ	0.143	0.14	0.143	0.142	0.139
average deviation over [0, 0.5]	0.000043	0.000105	0.000372	0.003559	0.014851

corresponding scale factor (which actually is $1/\alpha$), and in the last row the average squares-based deviation over the interval $[0, 0.5]$ is shown.

SCF7 Scaled sum with interaction terms

A sometimes used combination of sum and multiplication is obtained in the following format:

$$c(V_1, \dots, V_k) = (V_1 + \dots + V_k)/\lambda + \sum_{i,j} \mu_{i,j} V_i V_j$$

One example is given in terms of the sum function as follows:

$$c(V_1, \dots, V_k) = 4(V_1 + \dots + V_k)(1 - (V_1 + \dots + V_k))$$

This is an example of a combination function which is not monotonic. It has an U-shape with its maximum 1 when $V_1 + \dots + V_k = 0.5$, and for other values of $V_1 + \dots + V_k$ it is lower than 1.

SCF8 Advanced product-based combination function

In the combined product-based combination function, the two variants are combined by taking their weighted average with weight factors β and $1 - \beta$ for a parameter β with $0 \leq \beta \leq 1$. The overall combination function is obtained as

$$\begin{aligned} c(V_1, \dots, V_k) &= \mathbf{aproduct}_\beta(V_1, \dots, V_k) \\ &= (1 - \beta) \mathbf{product}(V_1, \dots, V_k) + \beta \mathbf{cproduct}(V_1, \dots, V_k) \end{aligned}$$

or

$$\beta(1 - (1 - V_1)_{*} \dots (1 - V_k)) + (1 - \beta)V_{1*} \dots V_k$$

Here β is a weight factor that can be used as *bias parameter*. The higher the value of β , the more upward bias, and the lower this value, the more downward bias.

For example:

$$c(V_1, V_2) = \mathbf{aproduct}_\beta(V_1, V_2) = \beta(1 - (1 - V_1)(1 - V_2)) + (1 - \beta)V_1 V_2$$

For this case of two arguments $\beta < 0.5$ indicates a downward bias and $\beta > 0.5$ an upward bias; for $\beta = 0.5$ a scaled sum function is obtained:

$$\begin{aligned}\mathbf{aproduct}_{0.5}(V_1, V_2) &= 0.5(1 - (1 - V_1)(1 - V_2)) + 0.5V_1V_2 \\ &= 0.5(V_1 + V_2 - V_1V_2) + 0.5V_1V_2 \\ &= (V_1 + V_2)/2 \\ &= \mathbf{ssum}_2(V_1, V_2)\end{aligned}$$

SCF9 Advanced min and max-based combination function

$$\begin{aligned}c(V_1, \dots, V_k) &= \mathbf{aminmax}_\beta(V_1, \dots, V_k) \\ &= (1 - \beta)\mathbf{min}(V_1, \dots, V_k) + \beta\mathbf{max}(V_1, \dots, V_k)\end{aligned}$$

Here again β is a *bias parameter*: as for the case of the advanced product function, the higher the value of β , the more upward bias, and the lower this value, the more downward bias.

SCF10 Advanced composed product and scaled sum combination function

$$\begin{aligned}c(V_0, \dots, V_k) &= \mathbf{aproduct} - \mathbf{ssum}_\beta(V_0, V_1, \dots, V_k) \\ &= \mathbf{aproduct}_\beta(V_0, \mathbf{ssum}(V_1, \dots, V_k))\end{aligned}$$

This is an example of another way in which standard combination functions can be used as building blocks, in this case by composing them (applying one to the result of the other). This can be (and actually has been) used, for example, in models for social contagion where first by a scaled sum combination function $\mathbf{ssum}(V_1, \dots, V_k)$ the average level of the levels $V_1, \dots, V_k$ of the other, connected persons is determined, and the result of this is combined with the own level V_0 by an advanced product function. In particular, this has been applied to model (amplified) emotion contagion spirals; see (Bosse et al. 2009, 2015). This function satisfies the same properties CFP1 to CFP5 from Sect. 2.7 as the scaled sum function, but does not satisfy the symmetry property SCF6 concerning its first argument.

2.7 Properties for Combination Functions

Combination functions can satisfy certain properties. A number of such properties are discussed here. In Table 2.10 it is shown which of these standard combination functions satisfy which of these properties.

CFP1 Preservation of zero impact

A combination function $c(\cdot)$ *preserves zero impact*, if it does never provide nonzero aggregated impact out of zero single impacts:

Table 2.10 Overview of standard combination functions and their properties

Name	Description	Formula	CFP1	CFP2	CFP3	CFP4	CFP5
		$c(V_1, \dots, V_k) =$	Preservation of zero impact	Non-negative	Upward bounded by 1	Monotonic	Keeps states within $[0, 1]$
SCF1 sum(.)	Sum	$V_1 + \dots + V_k$	Yes	If all ω_{ji} $i \geq 0$	Depends	Yes	If all $\omega_{ji} \geq 0$ & $\sum_j \omega_{ji} \leq 1$
SCF2 product(.)	Product	$V_1 * \dots * V_k$	Yes	If all ω_{ji} $i \geq 0$	If all $V_i \leq 1$	Yes	If all $\omega_{ji} \geq 0$
SCF3 min(.)	Complement product	$1 - (1 - V_1) * \dots * (1 - V_k)$	Yes	If all ω_{ji} $i \geq 0$	If all $V_i \leq 1$	Yes	If all $\omega_{ji} \geq 0$
SCF4 max(.)	Minimal value	$\min(V_1, \dots, V_k)$	Yes	If all ω_{ji} $i \geq 0$	If all $V_i \leq 1$	Yes	If all $\omega_{ji} \geq 0$
SCF5 alogistic _{σ, τ} (.)	Maximal value	$\max(V_1, \dots, V_k)$	Yes	If all ω_{ji} $i \geq 0$	If all $V_i \leq 1$	Yes	If all $\omega_{ji} \geq 0$
SCF6 ssum _{λ} (.)	Simple logistic sum	$1/(1 + e^{-\sigma(V_1 + \dots + V_k - \tau)})$ with $\sigma, \tau \geq 0$	No	Yes	Yes	Yes	Yes
SCF7 sisum(.)	Advanced logistic sum	$[1/(1 + e^{-\sigma(V_1 + \dots + V_k - \tau)}) - (1/(1 + e^{\sigma}))]/(1 + e^{-\sigma\tau})$ with $\sigma, \tau \geq 0$	Yes	Yes	Yes	Yes	If all $\omega_{ji} \geq 0$
SCF8 aproduct _{β} (.)	Scaled sum	$(V_1 + \dots + V_k)/\lambda$ with $\lambda > 0$	Yes	If all ω_{ji} $i \geq 0$	Depends	Yes	If all $\omega_{ji} \geq 0$ & $\lambda \geq \sum_j \omega_{ji}$
SCF9 aminmax _{β} (.)	Scaled sum with interaction terms	$(V_1 + \dots + V_k)/\lambda + \sum_{ij} \mu_{ij} V_i V_j$ with $\lambda > 0$	Yes	Depends	Depends	Depends	Depends
SCF10 aproduct-ssum _{β, λ} (.)	Advanced product	$\beta \text{ cproduct}(V_1, \dots, V_k)$ with $0 \leq \beta \leq 1$	Yes	Yes	Yes	Yes	If all $\omega_{ji} \geq 0$
SCF11 aminmax _{β} (.)	Advanced minimum and maximum	$\beta \text{ max}(V_1, \dots, V_k)$ with $0 \leq \beta \leq 1$	Yes	Yes	Yes	Yes	If all $\omega_{ji} \geq 0$
SCF12 aproduct-ssum _{β, λ} (.)	Advanced composed product and scaled sum	$\beta \text{ max}(V_1, \dots, V_k)$ with $0 \leq \beta \leq 1$ & $\lambda > 0$	Yes	Yes	Yes	Yes	If all $\omega_{ji} \geq 0$ & $\lambda \geq \sum_j \omega_{ji}$

$$c(0, \dots, 0) = 0$$

Almost all standard combination functions satisfy this property CFP1. The only exception above is the simple logistic sum combination function SCF4.

CFP2 Nonnegative

A combination function $c(\cdot)$ is *nonnegative* if for nonnegative single impacts it never provides negative aggregated impact:

$$V_i \geq 0 \text{ for all } i \Rightarrow c(V_1, \dots, V_k) \geq 0$$

Sometimes combination functions are defined in such a way that this automatically holds:

$$c^*(V_1, \dots, V_k) = \begin{cases} c(V_1, \dots, V_k) & \text{if } c(V_1, \dots, V_k) \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

CFP3 Upward bounded by 1

A combination function $c(\cdot)$ is *upward bounded by 1* if it always provides aggregated impact values not exceeding 1:

$$V_i \leq 1 \text{ for all } i \Rightarrow c(V_1, \dots, V_k) \leq 1$$

Sometimes combination functions are defined in such a way that this automatically holds:

$$c^*(V_1, \dots, V_k) = \begin{cases} c(V_1, \dots, V_k) & \text{if } c(V_1, \dots, V_k) \leq 1 \\ 1 & \text{otherwise} \end{cases}$$

Note that the sum combination function SCF1 does not satisfy this property, but the product combination function SCF2 does, as do the minimum and maximum, and the logistic sum combination functions SCF3 and SCF4. The advanced product-based combination function SCF8 also satisfies property CFP3, as can be seen (for two impacts) as follows.

If $V_i \leq 1$, it holds

$$\mathbf{product}(V_1, V_2) = V_1 V_2 \leq 1$$

and

$$\begin{aligned} (1 - V_1)(1 - V_2) &\geq 0 \\ -(1 - V_1)(1 - V_2) &\leq 0 \\ 1 - (1 - V_1)(1 - V_2) &\leq 1 \end{aligned}$$

So, also

$$\mathbf{cproduct}(V_1, V_2) \leq 1$$

Therefore

$$\begin{aligned} \mathbf{aproduct}(V_1, V_2) &= \beta \mathbf{cproduct}(V_1, V_2) + (1 - \beta) \mathbf{product}(V_1, V_2) \\ &= \beta(1 - (1 - V_1)(1 - V_2)) + (1 - \beta)V_1 V_2 \\ &\leq \beta + (1 - \beta) \\ &= 1 \end{aligned}$$

CFP4 Monotonicity

A combination function $c(\cdot)$ is *monotonic*, if higher single impacts result in a higher aggregated impact:

$$c(V_1, \dots, V_k) \leq c(W_1, \dots, W_k) \text{ whenever } V_i \leq W_i \text{ for all } i$$

All standard combination functions SCF1 to SCF4 above satisfy this property. However, sometimes combination functions may be relevant that are not monotonic. For example, outdoor circumstances such as temperature and wind speed may have impact on how happy you feel. Changing from extremely low to a bit higher these variables may have an increased impact on you being happy. But changing from there to extremely high temperature and wind speed may have a decreased impact on you being happy. So, for some combinations of causal relations it may hold that the higher the single impacts, the higher the aggregated impact (monotonic combination function), but for some other combinations of causal relations it may be different; for example, there may be single impact levels that generate a maximal aggregated impact, and both above and below these single impact levels the aggregated impact is lower. The example

$$c(V_1, \dots, V_k) = 4(V_1 + \dots + V_k)(1 - (V_1 + \dots + V_k))$$

of a combination function SCF7 given above models such a situation.

CFP5 Keeps values within [0, 1]

A combination function $c(\cdot)$ *keeps the values of the states X_i within the interval [0, 1]*, if for $\Delta t \leq 1$ and $\eta_i \leq 1$ it holds:

$$0 \leq X_i(t) \leq 1 \text{ for all } i \Rightarrow 0 \leq X_i(t + \Delta t) \leq 1 \text{ for all } i$$

This property is entailed by two other properties:

$$\mathbf{CFP2} \ \& \ \mathbf{CFP3} \Rightarrow \mathbf{CFP5}$$

Any nonnegative combination function which is upward bounded by 1 keeps any variable within the interval [0, 1], assuming that it starts in that interval. This

can be shown as follows. When at time t for all activation values it holds $0 \leq X_i(t) \leq 1$, then for $\Delta t \leq 1$ and $\eta_i \leq 1$, also $0 \leq X_i(t + \Delta t) \leq 1$; see also Fig. 2.5. Suppose that this combination function is used for a state X_i with $X_1, \dots, X_k$ connected toward X_i . Then

$$\begin{aligned} X_i(t + \Delta t) &= X_i(t) + \eta_i[c_i(\dots) - X_i(t)]\Delta t \\ &= X_i(t) (1 - \eta_i\Delta t) + \eta_i c_i(\dots)\Delta t \\ &\geq 0 \end{aligned}$$

Moreover,

$$\begin{aligned} X_i(t + \Delta t) &= X_i(t) (1 - \eta_i\Delta t) + \eta_i c_i(\dots)\Delta t \\ &\leq (1 - \eta_i\Delta t) + \eta_i\Delta t \\ &= 1 \end{aligned}$$

As this CFP5 is a convenient property for modeling, often combination functions are chosen that fulfill this property.

Finally, the following property is shared by all standard combination functions presented in Sect. 2.6, except for SCF10.

CFP6 Symmetric

A combination function $c(\dots)$ is *symmetric* if any permutation of its arguments provides the same result:

$$c(U_1, \dots, U_k) = c(V_1, \dots, V_k)$$

when $U_1, \dots, U_k$ is any permutation of $V_1, \dots, V_k$.

2.8 Applying Computational Methods to Model Representations

Multiple computational methods can be applied to a declarative conceptual or numerical representation as discussed in the above sections, varying from simulation methods to analysis methods. Examples of applying such computational methods are shown in Table 2.11.

As a way of applying a computational method to the declarative numerical description for this example model, these difference equations can be used calculate for all states (in parallel) for each time point t the values at time point $t + \Delta t$. For an example, see Table 2.13 (for time points 0–20), and see Fig. 2.21 for a graphical representation of the simulation results up to time point 60. In Table 2.13 each row represents the overall state S at the time point t indicated in the left column. This state is described by the values for the six different specific states or state properties distinguished in the model in each of the columns: from world state for stimulus s to

Table 2.11 Examples of computational methods that can be applied to a temporal-causal network model

Conceptual representation	Numerical representation
Qualitative causal reasoning	Forward simulation, backward simulation
Network analysis	Verification of equilibria, monotonicity
Conceptual validation	Numerical validation
Identification of qualitative personal or contextual characteristics in the context of empirical information	Identification of quantitative personal or contextual characteristics in the context of empirical information
Conceptual analysis or diagnostics of a process	Numerical analysis or diagnostics of a process
Conceptual planning of actions or interventions as part of some application	Numerical planning of actions or interventions as part of some application

Table 2.12 Matrix representation of the specific example model with simulation shown in Table 2.13 and Fig. 2.21

to from	ws_s	ss_s	srs_s	srs_e	ps_a	es_a
ws_s		1				
ss_s			1			
srs_s					0.5	
srs_e					0.5	
ps_a				1		1
es_a						
η_Y	-	0.8	0.8	0.8	0.8	0.4
$c_Y(\dots)$	-	id(.)	id(.)	id(.)	sum(.,.)	id(.)

execution state for action a . In Table 2.8 at each time point $t + \Delta t$ all the values of the corresponding row are calculated based on the values of the previous row for time point t (which in turn have been calculated in the previous time step).

In this example simulation all connection weights are 1, except that $\omega_{\text{responding}} = 0.5$ and $\omega_{\text{amplifying}} = 0.5$. Furthermore, $\Delta t = 0.5$ and the speed factors η_X for all states X are 0.8, except for the action execution es_a which is $\eta_{es_a} = 0.4$ (actions are slower than mental processes). The combination functions are as indicated above: all $c_X(\dots)$ are identity functions **id(.)** except the combination function $c_{ps_a}(\dots)$ for the preparation state ps_a , which is a sum function **sum(.,.)**. All these elements are indicated in the matrix representation in Table 2.12.

The table and graph show that at time 5 stimulus s occurs in the world (and disappears after time point 35): the icecream. Through the sensing connection this stimulus s is sensed from time point 5.5 on. Subsequently, through the representing

Table 2.13 Simulation values from time point 0 to time point 20 for the model depicted in Fig. 2.10 (time steps $\Delta t = 0.5$)

time	world state for stimulus s	sensor state for stimulus s	sensory representation of stimulus s	sensory representation of effect e	preparation state for action a	execution state for action a
0.0	0.00	0.00	0.00	0.00	0.00	0.00
0.5	0.00	0.00	0.00	0.00	0.00	0.00
1.0	0.00	0.00	0.00	0.00	0.00	0.00
1.5	0.00	0.00	0.00	0.00	0.00	0.00
2.0	0.00	0.00	0.00	0.00	0.00	0.00
2.5	0.00	0.00	0.00	0.00	0.00	0.00
3.0	0.00	0.00	0.00	0.00	0.00	0.00
3.5	0.00	0.00	0.00	0.00	0.00	0.00
4.0	0.00	0.00	0.00	0.00	0.00	0.00
4.5	0.00	0.00	0.00	0.00	0.00	0.00
5.0	1.00	0.00	0.00	0.00	0.00	0.00
5.5	1.00	0.40	0.00	0.00	0.00	0.00
6.0	1.00	0.64	0.16	0.00	0.00	0.00
6.5	1.00	0.78	0.35	0.00	0.03	0.00
7.0	1.00	0.87	0.52	0.01	0.09	0.01
7.5	1.00	0.92	0.66	0.04	0.16	0.02
8.0	1.00	0.95	0.77	0.09	0.24	0.05
8.5	1.00	0.97	0.84	0.15	0.31	0.09
9.0	1.00	0.98	0.89	0.22	0.39	0.13
9.5	1.00	0.99	0.93	0.28	0.45	0.18
10.0	1.00	0.99	0.95	0.35	0.52	0.24
10.5	1.00	1.00	0.97	0.42	0.57	0.29
11.0	1.00	1.00	0.98	0.48	0.62	0.35
11.5	1.00	1.00	0.99	0.53	0.66	0.40
12.0	1.00	1.00	0.99	0.59	0.70	0.46
12.5	1.00	1.00	0.99	0.63	0.74	0.50
13.0	1.00	1.00	1.00	0.67	0.77	0.55
13.5	1.00	1.00	1.00	0.71	0.79	0.59
14.0	1.00	1.00	1.00	0.75	0.82	0.63
14.5	1.00	1.00	1.00	0.77	0.84	0.67
15.0	1.00	1.00	1.00	0.80	0.86	0.71
15.5	1.00	1.00	1.00	0.82	0.88	0.74
16.0	1.00	1.00	1.00	0.84	0.89	0.76
16.5	1.00	1.00	1.00	0.86	0.90	0.79
17.0	1.00	1.00	1.00	0.88	0.91	0.81

connection from time point 6 on a sensory representation for s is developing, which by the responding connection leads to a preparation state for action a (to get the icecream) from time point 6.5 on. By the predicting connection the sensory representation of the action effect e becomes active (from time point 7 on), which in turn strengthens the activation of the preparation state for a through the amplifying connection (from time point 7.5 on). After that through the executing connection, from time point 8 on the action execution is developing stronger: getting the icecream.

As an illustration the simple model depicted in Fig. 2.10 was also simulated by using the advanced logistic function for two of the states: for the action preparation state ps_a and for the action execution state es_a . The for ps_a was set $\sigma_{ps_a} = 5$ and the

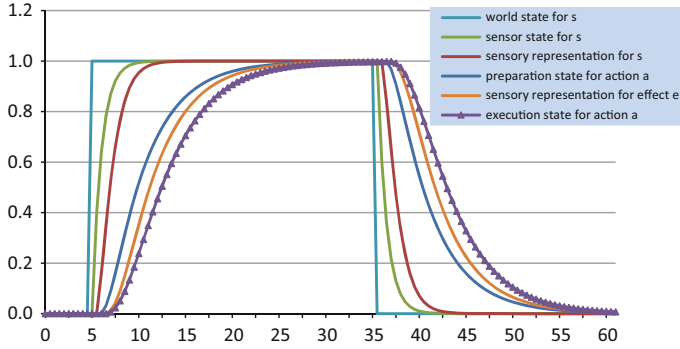


Fig. 2.21 Simulation example for the model depicted in Fig. 2.10 using identity and sum combination functions for all states

threshold $\tau_{ps_a} = 1$; the steepness for es_a was set $\tau_{es_a} = 20$ and the threshold $\tau_{es_a} = 0.6$; so:

$$c_{ps_a}(V_1, V_2) = \mathbf{alogistic}_{5,1}(V_1, V_2)$$

$$c_{es_a}(V) = \mathbf{alogistic}_{20,0.6}(V)$$

Moreover, this time all connection weights were made 1, also the connections to the preparation state ps_a . In Fig. 2.22 the results are shown. As can be seen the curves for the preparation and execution states for the action shown increases and decreases that are a bit more steep than in Fig. 2.21 where sum and identity combination functions were used for them.

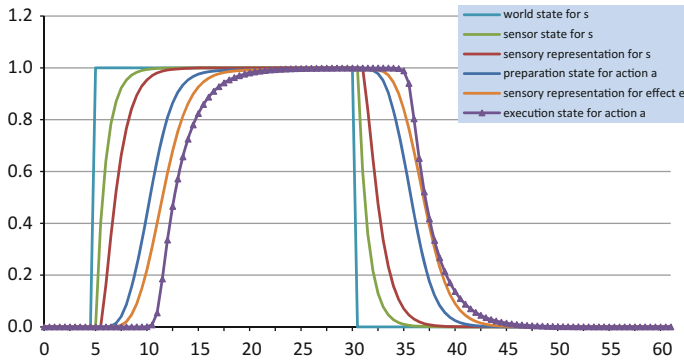


Fig. 2.22 Simulation example for the model depicted in Fig. 2.10 using advanced logistic sum combination functions for preparation and execution states

2.9 Applicability of the Modeling Perspective

The scope of applicability of the Network-Oriented Modeling approach covers mental processes described by causal networks of mental states, social (interaction) processes described by social network models, and more. In fact any scientific area in which causal relations are used to describe theories, hypotheses and findings falls within the scope of applicability. This covers practically all scientific domains, as causal explanation is used as a main vehicle almost everywhere in science. In this section first in Sect. 2.9.1 the notion of state-determined system is discussed in some more detail. Next, it is discussed in Sect. 2.9.2 how any smooth continuous state-determined system can be represented by a set of first-order differential equations and vice versa. Finally, in Sect. 2.9.3 it is discussed how any smooth continuous state-determined system (and therefore any model represented by a set of first-order differential equations) can be modeled by the temporal-causal network format described in Sects. 2.4 and 2.5.

2.9.1 The State-Determined System Assumption

Recall from Chap. 1, Sect. 1.5 that the notion of *state-determined system*, adopted from Ashby (1960) was taken as the basis to describe what a dynamical system is in van Gelder and Port 1995, p. 6. That a system is state-determined means that its *current state always determines a unique future behaviour*. This property is reflected in modeling and simulation. Three features in particular are (van Gelder and Port 1995):

- The future behaviour cannot depend on states the system might have been in *before* the current state: past history only can make a difference insofar as it has left an effect on the current state. This means that if you want to make a prediction on a next state, for example by simulation, only the information from the current state is needed, not from earlier states.
- That the current state determines future behaviour implies the existence of some *rule of evolution* describing the behaviour of the system as a function of its current state. The idea is that this rule can be specified in some reasonable succinct and useful fashion. The format introduced in Sects. 2.4 and 2.5 is an example of a format in which such rules of evolution can be expressed.
- That future behaviours are uniquely determined means that state space sequences can *never fork*. This means that when a next state is determined out of a current state, there is only one outcome.

The possibility of a choice of a proper set of state properties is the crucial factor to obtain a state-determined system that is practically usable. The validity of the assumptions underlying the Dynamical Systems Theory depends on the existence of such sets. For example, if to obtain a proper state-determined system to study some

mental process, all states of the universe (including, for example the positions of all planets and stars and even the mental states of all other humans) are needed, then for practical purposes this perspective is useless. The truth is that, even for those who believe in science, for example, concerning Newton's gravitation laws, even for the application of such solid monumental laws, in principle all mass and positions from the universe have to be taken into account, which obviously is infeasible. But in practice, only mass that is not very far away is incorporated in a model, which then in fact provides an approximation. If such an approximation is accurate enough (objects very far away have some effect, but this is a very small effect), then still a useful outcome can be obtained. So, more in general, usually an additional kind of locality assumption is made that to model a specific process (and not the whole universe), a limited set of state variables can be found to get a state-determined system. In Ashby (1960), such a hypothesis is expressed as follows:

Because of its importance, science searches persistently for the state-determined. As a working guide, the scientist has for some centuries followed the hypothesis that, given a set of variables, he can always find a larger set that (1) includes the given variables, and (2) is state-determined. Much research work consists of trying to identify such a larger set, for when it is too small, important variables will be left out of the account, and the behaviour of the set will be capricious. The assumption that such a larger set exists is implicit in almost all science, but, being fundamental, it is seldom mentioned explicitly (Ashby 1960, p. 28).

In this section it will be analyzed in some more depth in what formats in general state-determined systems can be described adequately. It will turn out that one such format is by sets of first-order differential equations, and another adequate format is the temporal-causal network format described in Sects. 2.4 and 2.5.

2.9.2 *State-Determined Systems and First-Order Differential Equations*

State-determined systems can be specified in mathematical formats; see (Ashby 1960, pp. 241–252) for some details. In the first place a finite set of states $X_1, \dots, X_n$ is assumed describing how the system changes over time via functions $X_1(t), \dots, X_n(t)$ of time t . The criteria for state-determined system can be formalized in a numerical manner by a relation (rule of evolution) that expresses that for each time point t the future value of each state X_i at time $t + s$ uniquely depends s and on $X_1(t), \dots, X_n(t)$ and hence can be described via some function $F_i(X_1, \dots, X_n, s)$ in the following manner (see also Ashby 1960, pp. 243–244):

$$X_i(t + s) = F_i(X_1(t), \dots, X_n(t), s) \text{ for } s \geq 0$$

Assuming continuous processes and smoothness (being differentiable) of the functions $X_i(t)$ and F_i , these relations can be reformulated (see Box 2.1) into a set of first order differential equations of the form

Suppose a smooth continuous state-determined system is given. A sketch of why it can be described by a set of first-order differential equations is as follows. For any given time point t the future states $X_i(t+s)$ at some future time point $t+s$ purely depend on s and the states $X_i(t)$ at t . This can be described by (smooth) mathematical functions $F_i(\dots)$:

$$X_i(t+s) = F_i(X_1(t), \dots, X_n(t), s) \text{ for } s \geq 0$$

In the particular case of $s = 0$ it holds

$$X_i(t) = F_i(X_1(t), \dots, X_n(t), 0)$$

Subtracting these two expressions above and dividing by s provides:

$$[X_i(t+s) - X_i(t)]/s = [F_i(X_1(t), \dots, X_n(t), s) - F_i(X_1(t), \dots, X_n(t), 0)]/s$$

When the limit for s very small, approaching 0 is taken, it follows that

$$dX_i(t)/dt = [\partial F_i(X_1(t), \dots, X_n(t), s)/\partial s]_{s=0}$$

Now define the function $f_i(X_1, \dots, X_n)$ by

$$f_i(X_1, \dots, X_n) = [\partial F_i(X_1, \dots, X_n, s)/\partial s]_{s=0}$$

Then it holds

$$dX_i(t)/dt = f_i(X_1(t), \dots, X_n(t))$$

This shows that the given state-determined system can be described by a set of first-order differential equations.

Box. 2.1 Why a smooth continuous state-determined system can be represented by a set of first-order differential equations

$$dX_i(t)/dt = f_i(X_1(t), \dots, X_n(t))$$

for some functions $f_i(X_1, \dots, X_n)$ (see also Ashby 1960, pp. 244–246). Note that X_i may also occur in $f_i(X_1, \dots, X_n)$.

Conversely, such a set of first-order differential equations always describes a state-determined system. So for the smooth continuous numerical case, state-determined systems are the systems that can be described by sets of first-order differential equations (Ashby 1960, p. 246).

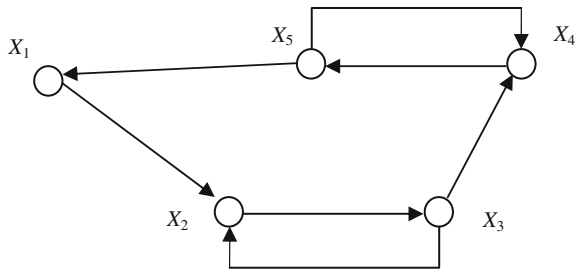
2.9.3 *State-Determined Systems and Modeling Based on Temporal-Causal Networks*

Sets of first order differential equations form a very general format used in computational modeling in many disciplines. However, also in many disciplines, processes are described and explained in terms of causal relationships. It would be helpful for understanding to relate these two perspectives in a transparent conceptual and mathematical manner. This will be discussed here.

For cognitive and neurological modeling in particular, often causal relationships are used in explaining mental processes (of the type partly depicted in Fig. 2.1). But also in many other domains, in a wide variety of scientific disciplines causal relationships play a crucial role. An example is the perspective called System Dynamics (e.g., Forrester 1973, 1987) in which at a basic conceptual level processes are described by means of graphs with states and arrows indicating causal relationships. In this context it will be useful if it can be explained more explicitly how any state-determined system can be described or transformed into a format that more directly relates to causal relationships between states. This indeed can always be achieved in the format of Sects. 2.4 and 2.5 (when arbitrary combination functions are allowed), in the manner shown in the following. Note that below in the current Sect. 2.9.3 (see Fig. 2.23), and also in Sect. 2.10 specific examples can be found of such a transformation.

In a state-determined system the changes in each state S depend on the other states. Those states R that actually play a role in this dependence relation form a subset D_S of the set of all states (in some special cases this subset may be the set of all states). The states R not in this subset D_S are those states for which never any state change of R has influence on a change of the state of S . The states in this

Fig. 2.23 Conceptual graph representation for the example model based on the given differential equation representation



subset D_S can be considered to cause the changes in the state S . Such causal effects of states on each other by causal relationships can be visualized in a graphical manner as a conceptual representation of a temporal-causal network model such as shown in Sect. 2.4, for example, in Fig. 2.3. Such a network model can be simply defined by the following criterion: for any state R and any state S there is a connection from R to S if and only if $R \in D_S$. This provides a conceptual representation of the state-determined system as a causal graph.

This can also be done at the level of the numerical representation on the basis of a set of differential equations representing the state-determined system. Suppose a differential equation for one of the states X_i is given of the form:

$$dX_i(t)/dt = f_i(X_1(t), \dots, X_n(t))$$

Then this function $f_i(X_1(t), \dots, X_n(t))$ will depend on a subset D_{X_i} of the set of states $\{X_1, \dots, X_n\}$. Note that X_i may occur in D_{X_i} . Usually this function f_i will be given as a formula in $X_1, \dots, X_n$; then this subset can be taken as the set of all states in $\{X_1, \dots, X_n\}$ that actually occur in this formula. Again, for any two states X_j and X_i with $j \neq i$ a causal connection from X_j to X_i can be defined by the criterion that $X_j \in D_{X_i}$. Moreover, by defining the function $h_i(X_1, \dots, X_n)$ by

$$h_i(X_1, \dots, X_n) = X_i + f_i(X_1, \dots, X_n)$$

the above differential equation for X_i always can be rewritten into a differential equation of the form

$$dX_i(t)/dt = [h_i(X_1(t), \dots, X_n(t)) - X_i(t)]$$

for some function $h_i(X_1(t), \dots, X_n(t))$. This form is a specific case (for $\eta_i = 1$) of a more general model of the form

$$dX_i(t)/dt = \eta_i [h_i(X_1(t), \dots, X_n(t)) - X_i(t)]$$

where the parameter η_i indicates a speed factor for state X_i . Note again that X_i may occur in $h_i(X_1, \dots, X_n)$.

The obtained causal network model can be generalized further by incorporating more structure by introducing as additional parameters specific nonzero weight values $\omega_{j,i}$ for the causal connections from X_j to X_i . In that case the function $h_i(X_1, \dots, X_n)$ can be considered a combination function $c_i(X_1, \dots, X_n)$, where for this case for the connection weights it holds $\omega_{j,i} = 1$. Then the format found above can be considered as a specific case (for $\omega_{j,i} = 1$) of the still more general model of the form

$$dX_i(t)/dt = \eta_i [c_i(\omega_{1,i}X_1(t), \dots, \omega_{n,i}X_n(t)) - X_i(t)]$$

So, having started with any arbitrary continuous, smooth state-determined system and its representation

$$dX_i(t)/dt = f_i(X_1(t), \dots, X_n(t))$$

in differential equation format, finally a numerical representation of a temporal-causal network model according to Sect. 2.5 was obtained in the form:

$$dX_i(t)/dt = \eta_i[c_i(\omega_{1,i}X_1(t), \dots, \omega_{n,i}X_n(t)) - X_i(t)]$$

with $c_i(V_1, \dots, V_n)$ a combination function, and η_i and $\omega_{i,i}$ parameters for a speed factor and connection weights. The original state-determined system description is a special case of this temporal-causal network model for settings $\eta_i = 1$ and $\omega_{i,i} = 1$.

A discrete variant of this is the following difference equation:

$$X_i(t + \Delta t) = X_i(t) + \eta_i[c_i(\omega_{1,i}X_1(t), \dots, \omega_{n,i}X_n(t)) - X_i(t)]\Delta t$$

These differential and difference equation formats can be transformed into a conceptual representation of temporal-causal network models in the manner shown in Sects. 2.4 and 2.5 above. This shows that any continuous smooth state-determined system can be described by a specific temporal-causal network model as defined in Sects. 2.4 and 2.5, as long as any type of combination function is allowed.

As an illustration, consider an arbitrary example of a model described in by a numerical first-order differential equations representation:

$$\begin{aligned} dX_1(t)/dt &= X_1(t) (X_5(t) - 1) \\ dX_2(t)/dt &= X_1(t) - X_2(t) + X_3(t) \\ dX_3(t)/dt &= X_2(t) (1 - X_3(t)) \\ dX_4(t)/dt &= X_3(t) - X_4(t)(1 - X_5(t)) \\ dX_5(t)/dt &= X_5(t) (1 - X_4(t)) \end{aligned}$$

To determine the conceptual representation for this temporal-causal network model, the five states X_1, X_2, X_3, X_4, X_5 are considered. From each of the equations by inspecting which states occur in the right hand side it can subsequently be determined that

X_5	affects	X_1
X_1 and X_3	affect	X_2
X_2	affects	X_3
X_3 and X_5	affect	X_4
X_4	affects	X_5

These causal connections can be represented in graphical form as shown in Fig. 2.23, and in the matrix representation as shown in Table 2.14.

Note that, when comparing, for example, the first differential equation to the format defined in Sect. 2.5, it can be written as

$$\begin{aligned} \mathbf{d}X_1(t)/\mathbf{d}t &= X_1(t)(X_5(t) - 1) \\ &= X_1(t)(X_5(t) - 1) + X_1(t) - X_1(t) \\ &= [X_1(t)(X_5(t) - 1) + X_1(t)] - X_1(t) \end{aligned}$$

Here the part $(X_1(t)(X_5(t) - 1) + X_1(t))$ can be considered the result of a combination function

$$c_{X_1}(V_1, V_5) = V_1 + V_1(V_5 - 1)$$

applied to $X_1(t)$ (for V_1) and $X_5(t)$ (for V_5). In a similar manner the following combination functions can be identified from the differential equations

$$\begin{aligned} c_{X_1}(V_1, V_5) &= V_1 + V_1(V_5 - 1) &= V_1 V_5 \\ c_{X_2}(V_1, V_2, V_3) &= V_2 + V_1 - V_2 + V_3 &= V_1 + V_3 \\ c_{X_3}(V_2, V_3) &= V_3 + V_2(1 - V_3) &= V_2 + V_3 - V_2 V_3 \\ c_{X_4}(V_3, V_4, V_5) &= V_4 + V_3 - V_4(1 - V_5) &= V_3 + V_4 V_5 \\ c_{X_5}(V_4, V_5) &= V_5 + V_5(1 - V_4) &= 2V_5 - V_4 V_5 \end{aligned}$$

Note that these combination functions are not symmetric.

So, in this case the combination functions are built from sum and product functions. Using these functions the differential equations become:

$$\begin{aligned} \mathbf{d}X_1(t)/\mathbf{d}t &= c_{X_1}(X_1(t), X_5(t)) - X_1(t) \\ \mathbf{d}X_2(t)/\mathbf{d}t &= c_{X_2}(X_1(t), X_2(t), X_3(t)) - X_2(t) \\ \mathbf{d}X_3(t)/\mathbf{d}t &= c_{X_3}(X_2(t), X_3(t)) - X_3(t) \\ \mathbf{d}X_4(t)/\mathbf{d}t &= c_{X_4}(X_3(t), X_4(t), X_5(t)) - X_4(t) \\ \mathbf{d}X_5(t)/\mathbf{d}t &= c_{X_5}(X_4(t), X_5(t)) - X_5(t) \end{aligned}$$

Table 2.14 Matrix representation for the example model based on the given differential equation representation

to from	X_1	X_2	X_3	X_4	X_5
X_1	1	1			
X_2		1	1		
X_3		1	1	1	
X_4				1	1
X_5	1			1	1
η_Y	1	1	1	1	1
$c_Y(\dots)$	$V_1 V_5$	$V_1 + V_3$	$V_2 + V_3 - V_2 V_3$	$V_3 + V_4 V_5$	$2V_5 - V_4 V_5$

Table 2.15 Matrix representation for the general model subsuming the model based on the given differential equations

to from	X_1	X_2	X_3	X_4	X_5
X_1	ω_{X_1,X_1}	ω_{X_1,X_2}			
X_2		ω_{X_2,X_2}	ω_{X_2,X_3}		
X_3		ω_{X_3,X_2}	ω_{X_3,X_3}	ω_{X_3,X_4}	
X_4				ω_{X_4,X_4}	ω_{X_4,X_5}
X_5	ω_{X_5,X_1}			ω_{X_5,X_4}	ω_{X_5,X_5}
η_V	η_{X_1}	η_{X_2}	η_{X_3}	η_{X_4}	η_{X_5}
$c_V(\dots)$	$c_{X_1}(V_1, V_3)$	$c_{X_2}(V_1, V_2, V_3)$	$c_{X_3}(V_2, V_3)$	$c_{X_4}(V_3, V_4, V_5)$	$c_{X_5}(V_4, V_5)$

This is the numerical representation of a temporal-causal network model as described in Sect. 2.5 with $\eta_{X_i} = 1$ for all i and $\omega_{X_i,X_j} = 1$ for all i and j with connected states. It turns out that the model described by the differential equations representation is a special case of the numerical representation of the more general temporal-causal model described by the following differential equations:

$$\begin{aligned}
dX_1(t)/dt &= \eta_{X_1} [c_{X_1}(\omega_{X_1,X_1}X_1(t), \omega_{X_5,X_1}X_5(t)) - X_1(t)] \\
dX_2(t)/dt &= \eta_{X_2} [c_{X_2}(\omega_{X_1,X_2}X_1(t), \omega_{X_2,X_2}X_2(t), \omega_{X_3,X_2}X_3(t)) - X_2(t)] \\
dX_3(t)/dt &= \eta_{X_3} [c_{X_3}(\omega_{X_2,X_3}X_2(t), \omega_{X_3,X_3}X_3(t)) - X_3(t)] \\
dX_4(t)/dt &= \eta_{X_4} [c_{X_4}(\omega_{X_3,X_4}X_3(t), \omega_{X_4,X_4}X_4(t), \omega_{X_5,X_4}X_5(t)) - X_4(t)] \\
dX_5(t)/dt &= \eta_{X_5} [c_{X_5}(\omega_{X_4,X_5}X_4(t), \omega_{X_5,X_5}X_5(t)) - X_5(t)]
\end{aligned}$$

In matrix representation this more general model representation is described in Table 2.15.

2.10 Modeling Adaptive Processes by Adaptive Temporal-Causal Networks

Dynamical systems usually are described in a format in which a number of parameters represent some of the characteristics of the context they describe. For example, in the temporal-causal network format described in Sects. 2.4 and 2.5 such parameters can take the form of the connection weights and speed factors, and perhaps parameters in the combination functions used such as threshold and steepness parameters in logistic sum functions. Also in the general differential equations format usually a number of such parameters are included in functions in the right hand side of these differential equations. In principle such parameters have specific constant values for a given scenario, and these values represent the

characteristics of persons, social interactions and other contextual aspects in such a scenario. By varying these constant values different scenarios can be obtained (in each of which these values can be different but do not change over time). Such scenarios can be used to explore different situations, for example, different personality types of humans.

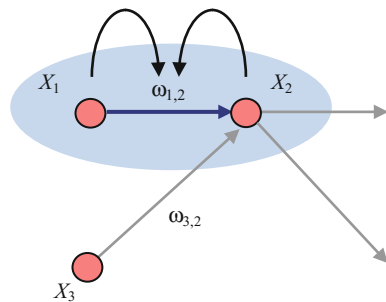
However, sometimes the modeled processes are adaptive in the sense that characteristics represented by such parameters can change over time as well, for example, as a result of learning. To deal with such adaptive behavior it is needed to interpret these adapting parameters as states that can change (i.e., as variables and no longer as parameters) and consider the resulting system as the dynamical system to be considered. To consider their dynamics, parameters conceptually have to be considered as states that change over time as well. This will be illustrated here for one specific case: the way in which connection strengths can change based on Hebbian learning. In Chap. 11 a similar type of adaptivity will be illustrated for adaptive social network models.

Hebbian learning (Hebb 1949), is based on the principle that strengthening of a connection between neurons over time may take place when both states are often active simultaneously ('neurons that fire together, wire together'); see also Fig. 2.24. The principle itself goes back to Hebb (1949), but over time has step by step gained more interest in the area of computational modeling due to more extensive empirical support (e.g., Bi and Poo 2001), and more advanced mathematical formulations (e.g., Gerstner and Kistler 2002).

More specifically, in the example model considered here it is assumed that the strength $\omega_{1,2}$ of such a connection between states X_1 and X_2 is adapted using the following *Hebbian learning rule*, taking into account a maximal connection strength 1, a *learning rate* $\eta > 0$, and an *extinction rate* $\zeta \geq 0$ (usually small), and activation levels $X_1(t)$ and $X_2(t)$ (between 0 and 1) of the two states involved. The first expression is in differential equation format, the second one in difference equation format

$$\begin{aligned} d\omega_{1,2}(t)/dt &= \eta X_1(t)X_2(t)(1 - \omega_{1,2}(t)) - \zeta \omega_{1,2}(t) \\ \omega_{1,2}(t + \Delta t) &= \omega_{1,2}(t) + [\eta X_1(t)X_2(t)(1 - \omega_{1,2}(t)) - \zeta \omega_{1,2}(t)]\Delta t \end{aligned}$$

Fig. 2.24 Hebbian learning principle



Such Hebbian learning rules can be found, for example, in Gerstner and Kistler (2002), p. 406. By the factor $1 - \omega_{1,2}$ the learning rule keeps the level of $\omega_{1,2}$ bounded by 1. When the extinction rate is relatively low, the upward changes during learning are proportional to both $X_1(t)$ and $X_2(t)$ and maximal learning takes place when both are 1. Whenever one of these activation levels is 0 (or close to 0) extinction takes over, and $\omega_{1,2}$ slowly decreases, if $\zeta > 0$. If $\zeta = 0$ no extinction takes place, then $\omega_{1,2}$ will never decrease.

It will be discussed how this can be modeled alternatively by considering the connection weight $\omega_{1,2}$ as a state that changes over time, represented by a node $\Omega_{1,2}$. Note that this is an example of the transformation discussed for the general case in Sect. 2.9.3.

As a first step an extra node for the state representing $\omega_{1,2}$ is added; see Fig. 2.25 and Table 2.16. This state, named $\Omega_{1,2}$, is affected by both X_1 and X_2 due to the learning, so connections from these states to $\Omega_{1,2}$ are needed. Moreover a connection from $\Omega_{1,2}$ to X_2 is needed to represent the effect of the connection strength on X_2 .

The weights of all of the connections from and to state $\Omega_{1,2}$ are assumed 1. As a next step it is explored what combination functions are needed for $\Omega_{1,2}$ and X_2 in this new situation depicted in Fig. 2.25. Suppose the previous situation depicted in Fig. 2.24 is described by the combination function $c_2(V_1, V_2)$ for X_2 which is applied to the impacts $\omega_{1,2}(t)X_1(t)$ and $\omega_{3,2}X_3(t)$ from X_1 and X_3 on X_2 of the form

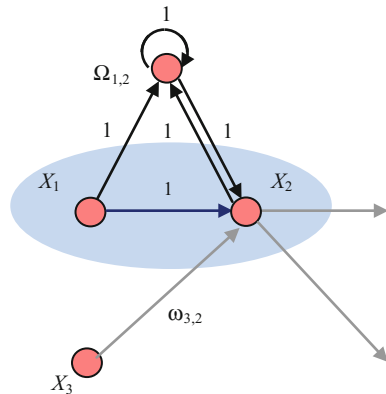
$$c_2(\omega_{1,2}(t)X_1(t), \omega_{3,2}X_3(t))$$

to obtain the difference equation for X_2

$$X_2(t + \Delta t) = X_2(t) + \eta_2 [c_2(\omega_{1,2}(t)X_1(t), \omega_{3,2}X_3(t)) - X_2(t)]\Delta t$$

In the new situation depicted in Fig. 2.25 the weight $\omega_{1,2}$ is represented by a state $\Omega_{1,2}$ which changes over time, with activation values $\Omega_{1,2}(t)$ the same as the

Fig. 2.25 Hebbian learning principle with a state $\Omega_{1,2}$ representing a dynamic connection weight



connection weight values $\omega_{1,2}(t)$ in the old situation for each t : $\Omega_{1,2}(t) = \omega_{1,2}(t)$. Now there are not two but three states with impact on X_2 , namely X_1 , X_3 and $\Omega_{1,2}$. This requires a combination function $c'_2(V_1, V_2, W)$ for X_2 with three arguments, which is applied to the impacts $X_1(t)$, $\omega_{3,2} X_3(t)$ and $\Omega_{1,2}(t)$ on X_2 , of the form

$$c'_2(X_2(t), \omega_{3,2} X_3(t), \Omega_{1,2}(t))$$

to obtain the difference equation for X_2

$$X_2(t + \Delta t) = X_2(t) + \eta_2 [c'_2(X_2(t), \omega_{3,2} X_3(t), \Omega_{1,2}(t)) - X_2(t)] \Delta t$$

This impact $c'_2(X_2(t), \omega_{3,2} X_3(t), \Omega_{1,2}(t))$ is the same as the impact $c_2(\omega_{1,2}(t) X_2(t), \omega_{3,2} X_3(t))$ in the previous model representation depicted in Fig. 2.24:

$$c'_2(X_2(t), \omega_{3,2} X_3(t), \Omega_{1,2}(t)) = c_2(\omega_{1,2}(t) X_2(t), \omega_{3,2} X_3(t))$$

So, recalling that $\Omega_{1,2}(t) = \omega_{1,2}(t)$ for all t , the new combination function can be defined as

$$c'_2(V_1, V_2, W) = c_2(W V_1, V_2)$$

For example, if $c_2(V_1, V_2)$ is the sum function $V_1 + V_2$, then

$$c'_2(V_1, V_2, W) = W V_1 + V_2$$

which is a combination of a product and a sum function. Alternatively, if, for example, $c_2(V_1, V_2)$ is the simple logistic sum function **slgistic**(V_1, V_2), then

$$c'_2(V_1, V_2, W) = \mathbf{slgistic}(W V_1, V_2) = 1 / (1 + e^{-\sigma(W V_1 + V_2 - \tau)})$$

Note that these combination functions $c'_2(V_1, V_2, W)$ are not symmetric.

Next, the combination function for the state $\Omega_{1,2}$ is identified that aggregates the impacts of X_1 and X_2 on $\omega_{1,2}$. The difference equation for the connection weight $\omega_{1,2}$ can be rewritten into one for the state $\Omega_{1,2}$ (with $\Omega_{1,2}(t) = \omega_{1,2}(t)$ for all t) as follows

$$\begin{aligned} \Omega_{1,2}(t + \Delta t) &= \Omega_{1,2}(t) + [\eta X_1(t) X_2(t) (1 - \Omega_{1,2}(t)) - \zeta \Omega_{1,2}(t)] \Delta t \\ &= \Omega_{1,2}(t) + [\eta X_1(t) X_2(t) (1 - \Omega_{1,2}(t)) - \zeta \Omega_{1,2}(t) + \Omega_{1,2}(t) - \Omega_{1,2}(t)] \Delta t \\ &= \Omega_{1,2}(t) + [\eta X_1(t) X_2(t) (1 - \Omega_{1,2}(t)) + (1 - \zeta) \Omega_{1,2}(t) - \Omega_{1,2}(t)] \Delta t \end{aligned}$$

From this it follows that the combination function $c_{\Omega_{1,2}}(V_1, V_2, W)$ for the state $\Omega_{1,2}$ should satisfy

$$c_{\Omega_{1,2}}(X_1(t), X_2(t), \Omega_{1,2}(t)) = \eta X_1(t) X_2(t) (1 - \Omega_{1,2}(t)) + (1 - \zeta) \Omega_{1,2}(t)$$

Table 2.16 Matrix representation for the Hebbian learning principle with a state $\Omega_{1,2}$

to from	X_1	X_2	X_3	$\Omega_{1,2}$
X_1		ω_{X_1, X_2}		$\omega_{X_1, \Omega_{1,2}}$
X_2				$\omega_{X_2, \Omega_{1,2}}$
X_3		ω_{X_3, X_2}		
$\Omega_{1,2}$		$\omega_{\Omega_{1,2}, X_2}$		$\omega_{\Omega_{1,2}, \Omega_{1,2}}$
η_Y	η_{X_1}	η_{X_2}	η_{X_3}	$\eta_{\Omega_{1,2}}$
$c_Y(\dots)$	$c_{X_1}(V_1, V_5)$	$c_{X_2}(V_1, V_2, V_3)$	$c_{X_3}(V_2, V_3)$	$c_{\Omega_{1,2}}(V_3, V_4, V_5)$

So, the speed factor $\eta_{\Omega_{1,2}}$ can be assumed 1 and the combination function for the description in Fig. 2.25 can be defined as:

$$c_{\Omega_{1,2}}(V_1, V_2, W) = \eta V_1 V_2 (1 - W) + (1 - \zeta) W$$

This function is a linear combination of sum and product functions. Note that this is an example of a combination function in which an impact of the state $\Omega_{1,2}$ on itself is included and aggregated; in the graph, this can be indicated by an arrow from $\Omega_{1,2}$ to itself. Moreover, note that this combination function is not symmetric and has two parameters η and ζ . In Table 2.17 the instantiated matrix representation of the generic one shown in Table 2.16 is shown.

As an alternative option (shown in Table 2.17), it is possible to assign weights to the two upward connections indicating links to the state $\Omega_{1,2}$: as a value $\sqrt{\eta}$ can be chosen or more general, η_1 and η_2 with $\eta = \eta_1 \eta_2$:

$$\begin{aligned} \Omega_{1,2}(t + \Delta t) &= \Omega_{1,2}(t) + [\eta_1 \eta_2 X_1(t) X_2(t) (1 - \Omega_{1,2}(t)) + (1 - \zeta) \Omega_{1,2}(t) - \Omega_{1,2}(t)] \Delta t \\ &= \Omega_{1,2}(t) + [(\eta_1 X_1(t)) (\eta_2 X_2(t)) (1 - \Omega_{1,2}(t)) + (1 - \zeta) \Omega_{1,2}(t) - \Omega_{1,2}(t)] \Delta t \end{aligned}$$

In this case:

$$c_{\Omega_{1,2}}(V_1, V_2, W) = V_1 V_2 (1 - W) + (1 - \zeta) W$$

This is shown in Table 2.18.

Table 2.17 Instantiated matrix representation for the Hebbian learning principle with a state $\Omega_{1,2}$

to from	X_1	X_2	X_3	$\Omega_{1,2}$
X_1		1		1
X_2				1
X_3		ω_{X_3, X_2}		
$\Omega_{1,2}$		1		1
η_Y		η_{X_2}		1
$c_Y(\dots)$		$W V_1 + V_2$		$\eta V_1 V_2 (1 - W) + (1 - \zeta) W$

Table 2.18 Instantiated matrix representation for the Hebbian learning principle with a state $\Omega_{1,2}$ with connections to $\Omega_{1,2}$ not 1

to from	X_1	X_2	X_3	$\Omega_{1,2}$
X_1		1		η_1
X_2				η_2
X_3		ω_{X_3, X_2}		
$\Omega_{1,2}$		1		1
η_Y		η_{X_2}		1
$c_Y(\dots)$		$W V_1 + V_2$		$V_1 V_2 (1 - W) + (1 - \zeta)W$

Yet another option is to also involve the extinction rate ζ as a weight $1 - \zeta$ of another connection from $\Omega_{1,2}$ to $\Omega_{1,2}$ itself:

$$\Omega_{1,2}(t + \Delta t) = \Omega_{1,2}(t) + [(\eta_1 X_1(t))(\eta_2 X_2(t)) (1 - \Omega_{1,2}(t)) + ((1 - \zeta)\Omega_{1,2}(t)) - \Omega_{1,2}(t)]\Delta t$$

with combination function for the four impacts on $\Omega_{1,2}$ of which the last two (indicated by the variables W_1 and W_2) are impacts from $\Omega_{1,2}$ itself (without parameters this time):

$$c_{\Omega_{1,2}}(V_1, V_2, W_1, W_2) = V_1 V_2 (1 - W_1) + W_2$$

This expresses that the state $\Omega_{1,2}$ is assumed to have some persistence, relating to the extinction rate ζ . Once it has some level, in principle it will keep this level, only extinction can make it (slowly) decrease. This can be modeled by assuming an additional connection with weight $1 - \zeta$ from the state $\Omega_{1,2}$ to itself, in addition to the connection from $\Omega_{1,2}$ to itself (with weight 1, shown in Fig. 2.25); see Fig. 2.26 and Table 2.19 for the new variant.

A slightly more sophisticated approach takes into account that at least some activation of state X_1 should occur prior to activation of X_2 . This can be modeled by adding an extra state X_4 that becomes active upon activation of X_1 and which is needed for activation of $\Omega_{1,2}$: it can be considered to indicate openness for learning; see Fig. 2.27.

In this section Hebbian learning was used as a basis for a case study of an adaptive network model. It was shown how as an alternative way of modeling the dynamic connection weights themselves can become states in the temporal-causal network model. In Chap. 11 it will be discussed how in similar manners adaptive social network models can be used to model social interaction.

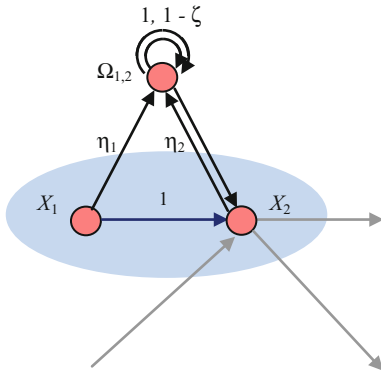


Fig. 2.26 Hebbian learning principle with a state $\Omega_{1,2}$ representing a connection weight and two connections to itself

Table 2.19 Instantiated matrix representation for the Hebbian learning principle with a state $\Omega_{1,2}$ with double connections from $\Omega_{1,2}$ to itself

to from	X_1	X_2	X_3	$\Omega_{1,2}$
X_1		1		η_1
X_2				η_2
X_3		ω_{X_3, X_2}		
$\Omega_{1,2}$		1		$\begin{matrix} 1 \\ 1-\zeta \end{matrix}$
η_Y		η_{X_2}		1
$c_Y(\dots)$		$W V_1 + V_2$		$V_1 V_2 (1 - W_1) + W_2$

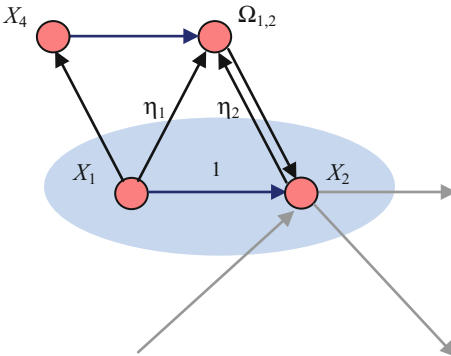


Fig. 2.27 Hebbian learning principle with an extra state for prior openness for learning

2.11 Discussion

The dynamic Network-Oriented Modeling approach based on temporal-causal networks discussed in this chapter and also presented earlier in Treur (2016) enables to design complex high level conceptual representations of temporal-causal network models, which can be systematically or automatically transformed into executable numerical model representations. The models are declarative; computational methods for simulation or analysis can be applied to them, but are independent of the model descriptions themselves. Dedicated software is available to support designing models in a conceptual manner (in graphical or matrix format). Moreover, using this software such conceptual representations can automatically be transformed into numerical representations in an executable format as a basis for performing simulation experiments.

In this chapter also the applicability has been discussed. In applications in many disciplines often state-determined systems are considered; e.g., (Ashby 1960; van Gelder and Port 1995). Following (Ashby 1960) it has been discussed how any set of first-order differential equations can be represented as a smooth state-determined system and conversely. Moreover, it was discussed how the temporal-causal network modeling approach can model any smooth state-determined system. This shows that the wide variety of applications based on state-determined systems (or first-order differential equations) all are covered.

The Network-Oriented Modeling approach based on temporal-causal networks used makes it easy to take into account theories and findings about dynamics of processes from any scientific discipline, as commonly such processes are described in terms of causal relations. In particular, this applies to complex brain processes known from Cognitive, Affective and Social Neuroscience, which often involve complex dynamics based on interrelating cycles. It enables to address in an integrative manner complex cognitive, affective and social phenomena such as dynamics related to social interactions, the integration of emotions within cognitive processes, internal simulation of external processes, mirroring of mental processes of others, and Hebbian learning; e.g., (Hebb 1949, Gerstner and Kistler 2002; Keysers and Perrett 2004; Keysers and Gazzola 2014).

It has been discussed how the approach relates to perspectives in Philosophy of Mind (e.g., Kim 1996), in particular to the notion of causal or functional role of a mental state, and how based on this notion networks of mental states can be used to model mental processes (see also Chap. 1). Furthermore, it has been discussed in Chap. 1 how such an approach relates to the philosophical perspective on dynamics in the world that is indicated as the clockwork universe; e.g., (Descartes 1634; Laplace 1825). This perspective relates to the notion of state-determined system; e.g., (Ashby 1960). The approach has been applied in a variety of domains, a number of which can be found in subsequent chapters.

References

- W.R. Ashby, *Design for a Brain* (Chapman and Hall, London (second extended edition, 1960), First edition, 1952)
- R.D. Beer, On the dynamics of small continuous-time recurrent neural networks. *Adapt. Behav.* **3**, 469–509 (1995)
- A. Bell, Levels and loops: the future of artificial intelligence and neuroscience. *Phil. Trans. R. Soc. Lond. B* **354**, 2013–2020 (1999)
- C. Bernard, *Introduction a l'etude de la medecine experimentale* (J. Bailliere et fils, Paris, 1865)
- G. Bi, M. Poo, M. Synaptic modification by correlated activity: Hebb's postulate revisited. *Annu. Rev. Neurosci.* **24**, 139–166 (2001)
- J. Bickle, *Psychoneural Reduction: The New Wave* (MIT Press, Cambridge, Mass, 1998)
- G. Booch, *Object Oriented Design with Applications* (Benjamins Cummins Publishing Company, Redwood City, 1991)
- T. Bosse, R. Duell, Z.A. Memon, J. Treur, C.N. van der Wal, A multi-agent model for emotion contagion spirals integrated within a supporting ambient agent model, in *Proceedings of the 12th International Conference on Principles of Practice in Multi-Agent Systems, PRIMA'09*, ed. by J.-J. Yang, M. Yokoo, T. Ito, Z. Jin, P. Scerri, Lecture Notes in Artificial Intelligence, vol. 5925 (Springer, 2009), pp. 48–67
- T. Bosse, R. Duell, Z.A. Memon, J. Treur, C.N. van der Wal, Agent-based modelling of emotion contagion in groups. *Cogn. Comput. J.* **7**(2015), 111–136 (2015)
- T. Bosse, J. Treur, Patterns in world dynamics indicating agency. *Trans. Comput. Collect. Intell.* **3** (2011), 128–151 (2011)
- B. Brewer, Self-location and agency. *Mind* **101**, 17–34 (1992)
- A. Damasio, *The Feeling of What Happens: Body, Emotion and the Making of Consciousness* (MIT Press, Cambridge, 1999)
- C. Dobbyn, S. Stuart, The self as an embedded agent. *Mind. Mach.* **13**, 187–201 (2003)
- D. Dubois, J. Lang, H. Prade, *Fuzzy sets in approximate reasoning*, part 2: logical approaches, 1991. *Fuzzy Sets Syst.* **40**, 203–244 (North-Holland, 1991)
- D. Dubois, H. Prade, Possibility theory, probability theory and multiple-valued logics: a clarification. *Ann. Math. Artif. Intell.* **32**, 35–66 (2002)
- J.W. Forrester, *World Dynamics*, (2nd ed.). Waltham, MA: Pegasus Communications (1973)
- J.W. Forrester, Lessons from System Dynamics Modeling. *System Dynamics Review* **3**, 136–149 (1987)
- K. Funahashi, Y. Nakamura, Approximation of dynamical systems by continuous time recurrent neural networks. *Neural Netw.* **6**, 801–806 (1993)
- W. Gerstner, W.M. Kistler, Mathematical formulations of Hebbian learning. *Biol. Cybern.* **87**, 404–415 (2002)
- G. Giangiacomo, *Fuzzy Logic: Mathematical Tools for Approximate Reasoning* (Kluwer Academic Publishers, Dordrecht, 2001)
- S. Grossberg, On learning and energy-entropy dependence in recurrent and nonrecurrent signed networks. *J. Stat. Phys.* **1**, 319–350 (1969)
- D. Hebb, *The Organisation of Behavior* (Wiley, 1949)
- M. Hirsch, Convergent activation dynamics in continuous-time networks. *Neural Networks* **2**, 331–349 (1989)
- J.J. Hopfield, Neural networks and physical systems with emergent collective computational properties. *Proc. Nat. Acad. Sci. (USA)* **79**, 2554–2558 (1982)
- J.J. Hopfield, Neurons with graded response have collective computational properties like those of two-state neurons. *Proc. Nat. Acad. Sci. (USA)* **81**, 3088–3092 (1984)
- C.M. Jonker, J.L. Snoep, J. Treur, H.V. Westerhoff, W.C.A. Wijngaards, BDI-modelling of complex intracellular dynamics. *J. Theor. Biol.* **251**(2008), 1–23 (2008)
- C.M. Jonker, J.L. Snoep, J. Treur, H.V. Westerhoff, W.C.A. Wijngaards, Putting intentions into cell biochemistry: an artificial intelligence perspective. *J. Theor. Biol.* **214**, 105–134 (2002a)

- C.M. Jonker, J. Treur, W.C.A. Wijngaards, Reductionist and antireductionist perspectives on dynamics. *Philos. Psychol. J.* **15**(2002), 381–409 (2002b)
- J.A.S. Kelso, *Dynamic patterns: the self-organization of brain and behavior* (MIT Press, Cambridge, 1995)
- C. Keysers, D.I. Perrett, Demystifying social cognition: a Hebbian perspective. *Trends Cogn. Sci.* **8**(2004), 501–507 (2004)
- C. Keysers, V. Gazzola, Hebbian learning and predictive mirror neurons for actions, sensations and emotions. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **369**, 20130175 (2014)
- J. Kim, *Philosophy of Mind* (Westview Press, 1996)
- D.E. Knuth, *The Art of Computer Programming* (Addison-Wesley, 1981)
- G. Kontonya, I. Sommerville, *Requirements Engineering: Processes and Techniques* (Wiley, New York, 1998)
- B.J. Kuipers, Commonsense reasoning about causality: deriving behavior from structure. *Artif. Intell.* **24**(1984), 169–203 (1984)
- B.J. Kuipers, J.P. Kassirer, How to discover a knowledge representation for causal reasoning by studying an expert physician, in ed. by F.R.G. Karlsruhe, *Proceedings Eighth International Joint Conference on Artificial Intelligence, IJCAI'83* (William Kaufman, Los Altos, CA, 1983)
- J. Pearl, *Causality*. Cambridge University Press, (2000)
- R.F. Port, T. van Gelder, *Mind as Motion: Explorations in the Dynamics of Cognition* (MIT Press, Cambridge, MA, 1995)
- S.M. Potter, What can artificial intelligence get from neuroscience?, in *Artificial Intelligence Festschrift: The next 50 years*, ed. by M. Lungarella, J. Bongard, R. Pfeifer (Springer, Berlin, 2007)
- K.R. Scherer, Emotions are emergent processes: they require a dynamic computational architecture. *Phil. Trans. R. Soc. B* **364**, 3459–3474 (2009)
- A. Sharpanskykh, J. Treur, Modelling and analysis of social contagion in dynamic networks. *Neurocomput. J.* **146**(2014), 140–150 (2014)
- A.S. Tanenbaum, *Structured Computer Organisation* (Prentice-Hall, London, 1976)
- J. Treur, Dynamic modeling based on a temporal-causal network modeling approach. *Biol. Insp. Cogn. Arch.* **16**, 131–168 (2016)
- T. van Gelder, R.F. Port, It's about time: an overview of the dynamical approach to cognition, in *Mind as motion: Explorations in the dynamics of cognition*, ed. by R.F. Port, T. van Gelder (Cambridge, MA: MIT Press), pp. 1–43
- L. Zadeh, Fuzzy sets as the basis for a theory of possibility. *Fuzzy Sets Syst.* **1**, 3–28, (1978). (Reprinted in *Fuzzy Sets and Systems* 100 (Supplement): 9–34, 1999)

Part II

Emotions All the Way

Chapter 3

How Emotions Come in Between Everything

Emotions Serving as Glue in All Mental and Social Processes

Abstract Within Cognitive, Affective and Social Neuroscience more and more mechanisms are found that suggest how emotions relate in a bidirectional manner to many other mental processes and behaviour. Based on this, in this chapter it is explored how a Network-Oriented Modeling approach can be used to model the dynamics and interaction of emotions. Using this approach it is illustrated by temporal-causal network models, for example, how emotions relate in a reciprocal manner to feelings, beliefs, desires, experiences, decision making, and to emotions of others.

3.1 Introduction

From the beginning Artificial Intelligence has addressed the modeling of cognitive processes behind intelligence. The original practice was that emotions were not taken into account in such models (for example, see Feigenbaum 1969). Presumably they were avoided since intelligence was aimed to be modeled in an idealised manner, and emotions were assumed to disturb that ideal. However, even in that time from the cognitive area it was pointed out that this was a serious omission when human intelligence is aimed at. For example, Neisser (1963) and Simon (1967) formulate this as follows:

Needs and emotions do not merely set the stage for cognitive activity and then retire. They continue to operate throughout the course of development. Moreover, they do not remain constant but undergo growth and change of their own, with substantial effects on intellectual activities (Neisser 1963, p. 196).

Information processing theories, however, have generally been silent on the interaction of cognition with affect. Since in actual human behavior motive and emotion are major influences on the course of cognitive behavior, a general theory of thinking and problem solving must incorporate such influences (Simon 1967, p. 29).

This situation in Artificial Intelligence has substantially changed in recent years. Currently, for conferences on AI, often modeling of emotions is one of the topics mentioned in their calls. One of the reasons for this change is the nowadays more

recognized need for human-like models, for example as a basis for virtual agents, or in Ambient Intelligence applications. Another reason for this change is the growing awareness fed by the strong development of neuroscience that in human-like models emotions cannot be neglected, as they play a role in most human processes, and this role often provides a constructive, and not a disturbing contribution. This widened scope of AI provides a multitude of new types of research questions that can be explored using computational modeling methods. Examples of such questions are:

- Does a feeling affect an expressed emotion or the other way around?
- In which way is it possible to control emotion?
- How does desiring relate to feeling?
- In how far do sensing and believing relate to feeling?
- How does having experiences over time relate to experiencing emotions?
- Can you make an adequate decision without feeling good about it?
- In how far is an individual in a group free in having own emotions?
- Why do groups with individuals with initially different preferences often come to common decisions and all members feel good with these decisions?

As discussed in Chap. 1, Sects. 1.2 and 1.3, structures and mechanisms found in neuroscience suggest that cyclic connections play an important role in many of the brain's processes. This holds in particular for the way in which emotions play their role. Emotions and feelings can be considered as being part of a number of inter-related adaptive and regulatory cycles, and based on these cycles emotional states emerge over time, and affect many other human processes. Examples of such types of cycles are emotional response—feeling cycles (e.g., Damasio 1999, 2010), emotion regulation cycles (e.g., Gross 1998; Goldin et al. 2008), cognitive-affective cycles (e.g., Phelps 2006; Pessoa 2008), and social contagion cycles (e.g., Iacoboni 2008; Hatfield et al. 2009). More advanced models for emotions and their role in mental functioning may involve a multiple of such types of cycles, that have to be integrated. One example of this further integration is described in Aziz et al. (2011) where an emotion regulation cycle is integrated with a social interaction cycle. As another example, in Hoogendoorn et al. (2011) and Bosse et al. (2012a) cognitive-affective cycles are integrated with social interaction cycles.

In the chapter, in subsequent sections, for a number of processes in which emotions play an important role it is discussed how the Network-Oriented Modeling perspective based on temporal-causal networks described in Chap. 2 can be applied. To get the idea of designing models according to this Network-Oriented Modeling approach, the emphasis in this chapter is on the design of the conceptual representations of these models, and on the generation of the numerical representations from these conceptual representations, and not on performing simulation experiments with them. The latter will be addressed in subsequent chapters for some specific models involving emotions. The chapter provides a unifying survey of a perspective for which different instances can be found in specific applications in the literature.

First, in Sect. 3.2 the cycle between emotional response and feeling is addressed, after which in Sect. 3.3 emotion regulation is discussed. Furthermore, in Sect. 3.4 the interaction between cognitive and affective states is discussed, and in Sect. 3.5 the role of emotion-related valuing in decision making. Finally, emotions in social contagion processes are addressed in Sect. 3.6. The chapter closes with a discussion.

3.2 Generating Emotional Responses and Feelings

The question on the direction of causality between feeling and emotional response has a long history. A classical view on emotions is that based on some sensory input, due to internal processing emotions are felt, and based on this they are expressed in some emotional response (e.g., a body state such as a face expression):

stimulus → sensory representation → felt emotion → preparation for bodily changes → expressed emotion

James (1884) claimed a different direction of causality (see also Damasio 2010, pp. 114–116):

stimulus → sensory representation → preparation for bodily changes → expressed emotion
→ felt emotion

The perspective of James assumes that a *body loop* via the expressed emotion is used to generate a felt emotion by sensing the own body state. Damasio made a further step by introducing the possibility of an *as-if body loop* bypassing actually expressed bodily changes (Damasio 1994, pp. 155–158; see also Damasio 1999, pp. 79–80; Damasio 2010):

stimulus → sensory representation → preparation for bodily changes → felt emotion

An as-if body loop describes an *internal simulation* of the bodily processes, without actually affecting the body, comparable to simulation in order to perform, for example, prediction, mindreading or imagination; e.g., Becker and Fuchs (1985), Goldman (2006) and Hesslow (2002). Damasio (1999, 2010) distinguishes an emotion (or emotional response) from a feeling (or felt emotion). A brief survey of Damasio's ideas about emotion and feeling can be found in Damasio (2010, pp. 108–129). According to this perspective emotions relate to actions, whereas feelings relate to perceptions of own body states:

Emotion and feeling, albeit part of a tightly bound cycle, are distinguishable processes. (...) Emotions are complex, largely automated programs of *actions* concocted by evolution. The actions are complemented by a *cognitive* program that includes certain ideas and modes of cognition, but the world of emotions is largely one of actions carried out in our bodies, from facial expressions and postures to changes in viscera and internal milieu. Feelings of emotion, on the other hand, are composite *perceptions* of what happens in our body and mind when we are emoting. As far as the body is concerned, feelings are images of actions rather than actions themselves; the world of feelings is one of perceptions executed in brain maps. (...) While emotions are actions accompanied by ideas and certain modes of

thinking, emotional feelings are mostly perceptions of what our bodies do during the emoting, along with perceptions of our state of mind during that same period of time (Damasio 2010, pp. 109–110).

Seen from a neural perspective, the emotion-feeling cycle begins in the brain, with the perception and appraisal of a stimulus potentially capable of causing an emotion and the subsequent triggering of an emotion. The process then spreads elsewhere in the brain and in the body proper, building up the emotional state. In closing, the process returns to the brain for the feeling part of the cycle, although the return involves brain regions different from those in which it all started (Damasio 2010, p. 111).

The emotion and feeling in principle mutually affect each other in a bidirectional manner: an as-if body loop usually occurs in a cyclic form by assuming that the emotion felt in turn affects the prepared bodily changes; see, for example, in Damasio (2010, pp. 119–122):

emotion felt → preparation for bodily changes

A brief survey of Damasio's ideas about emotion and feeling, and the 'tightly bound cycle' between them can be found in Damasio (2003, pp. 91–92) and Damasio (2010, pp. 108–129); for example:

The brain has a direct means to respond to the object as feelings unfold because the object at the origin is inside the body, rather than external to it. The brain can act directly on the very object it is perceiving. It can do so by modifying the state of the object, or by altering the transmission of signals from it. The object at the origin on the one hand, and the brain map of that object on the other, can influence each other in a sort of reverberative process that is not to be found, for example, in the perception of an external object. (...)

In other words, feelings are not a passive perception or a flash in time, especially not in the case of feelings of joy and sorrow. For a while after an occasion of such feelings begins – for seconds or for minutes – there is a dynamic engagement of the body, almost certainly in a repeated fashion, and a subsequent dynamic variation of the perception. We perceive a series of transitions. We sense an interplay, a give and take (Damasio 2003, pp. 91–92).

So, an as-if body loop usually occurs in a cyclic form by assuming that the emotion felt in turn affects the prepared bodily changes (see also Fig. 3.1, lower part):

emotion felt = based on sensory representation of (simulated) body state b → preparation for body state b = emotional response

For emotions Damasio describes the following biological substrate:

Emotions work when images processed in the brain call into action a number of emotion-triggering regions, for example, the amygdala or special regions of the frontal lobe cortex. Once any of these trigger regions is activated, certain consequences ensue – chemical molecules are secreted by endocrine glands and by subcortical nuclei and delivered to both the brain and the body (e.g., cortisol in the case of fear), certain actions are taken (e.g., fleeing or freezing; contraction of the gut, again in the case of fear), and certain expressions are assumed (e.g., a face and posture of terror) (Damasio 2010, p. 110).

Note that here a role of the amygdala is indicated in the process of generating an emotion, whereas in earlier times often the amygdala was related to feelings. In contrast, Damasio describes the substrate for feelings as follows:

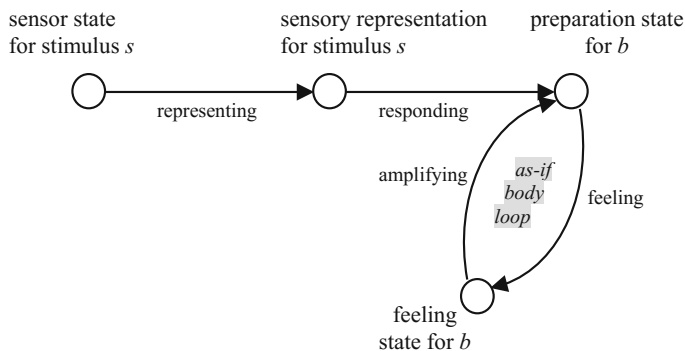


Fig. 3.1 Graphical conceptual representation of a temporal-causal network model for generating emotions and feelings based on a cyclic as-if body loop

In the late 1980s I hypothesized a role for the somatosensory cortices in feelings, and I pointed to the insula as a likely provider of feelings. I wanted to move away from the hopeless idea of attributing the origin of feeling states to action-driving regions, such as the amygdalae (Damasio 2010, p. 118).

At that time this idea had a rather hypothetical character, and was not the accepted view. This changed after 2000:

Since 2000, however, we have known that activity in the insula is indeed an important correlate for every conceivable kind of feeling (...) The idea that the insular cortex is an important substrate for feelings is certainly correct. (...) The anterior cingulate cortex tends to become active in parallel with the insula when we experience feelings. The insula and anterior cingulate are closely interlocked regions, the two being joined by multiple connections. The insula has dual sensory and motor functions, albeit biased toward the sensory side of the process, while the anterior cingulate operates as a motor structure (Damasio 2010, p. 118).

In addition to these, the process of generating a feeling involves several sub-cortical regions for certain preprocessing as well, as ‘they are the first recipients of information from the viscera and internal milieu with the ability to integrate signals from the entire range of the body’s interior’ (Damasio 2010, pp. 118–119).

This essentially shows a cyclic process that (for a constant environment) can lead to equilibrium states for both emotional response (preparation) and feeling. These biological mechanisms as briefly sketched have been used to obtain a conceptual representation of a temporal-causal network model depicted as a graph in Fig. 3.1. Here b is a label indicating a specific body state corresponding to the considered emotion. Note that what is called stimulus s here can be taken as the sensor state sensing s . This answers the question on the direction of the causality between feeling and emotional response in the sense that both emotional response affects feeling and feeling affects emotional response, in a cyclic manner. Note that for stimuli s and body states b indicate abstract states which by themselves may be characterised by multiple aspects; see also, for example, Damasio (1999), Lazarus (1991), Roseman (1996), Scherer (1999, 2009). For example preparation for a

Table 3.1 Conceptual matrix representation for the temporal-causal network model described in Fig. 3.1

To From	ss_s	srs_s	ps_b	fs_b
ss_s		$\omega_{\text{representing}}$		
srs_s			$\omega_{\text{responding}}$	
ps_b				ω_{feeling}
fs_b			$\omega_{\text{amplifying}}$	
η_Y	—	η_{srs_s}	η_{ps_b}	η_{fs_b}
$c_Y(\dots)$	—	$c_{srs_s}(V)$	$c_{srs_s}(V_1, V_2)$	$c_{fs_b}(V)$

specific emotional response b can involve different aspects of the body such as heart rate, skin, and specific chemicals in the blood. In many cases these abstract states can be related to vectors of values for such multiple aspects (see also Bosse et al. 2008a). Moreover, in more complex models more than one stimulus s and more than one body state b can be modeled as abstract states, for example, as $s_1, s_2, \dots$ and $b_1, b_2, \dots$. Then for each combination of an s_i and a b_j relations as depicted in Fig. 3.1 can be covered by the model (multiple stimuli have a combined effect on each preparation state). Note that the perspective presented here uses explicitly represented feeling states, which contrasts to approaches that consider emotions to be not explicitly represented as states having a causal effect; see, for example, Peck and Kozloski (2011).

In matrix representation the model shown in Fig. 3.1 is described by Table 3.1.

The conceptual representation of a temporal-causal network model described in Fig. 3.1 has been used to obtain a numerical representation of this temporal-causal network model, according to the systematic approach described in Chap. 2; see Box 3.1. Here each of the states has an associated differential equation for the update of its activation level, and all connection weights are nonnegative. Note that due to monotonicity of the combination functions $c_X(\dots)$, it holds that the higher the value of ω_{feeling} , the stronger the level of the response state affects the level of the feeling state fs_b , and the higher the value of $\omega_{\text{amplifying}}$, the stronger the feeling level affects the emotional response level ps_b . Due to the cyclic nature of the model this effect propagates back and forth between the two states.

LP1 Representing stimulus s

$$dsrs_s/dt = \eta_{srs_s} [c_{srs_s}(\omega_{\text{representing}} ss_s) - srs_s]$$

$$srs_s(t + \Delta t) = srs_s(t) + \eta_{srs_s} [c_{srs_s}(\omega_{\text{representing}} ss_s(t)) - srs_s(t)] \Delta t$$

LP2 Generating and amplifying preparation for response b

$$dps_b/dt = \eta_{ps_b} [c_{ps_b}(\omega_{\text{responding}} srs_s, \omega_{\text{amplifying}} fs_b) - ps_b]$$

$$ps_b(t + \Delta t) = ps_b(t) + \eta_{ps_b} [c_{ps_b}(\omega_{\text{responding}} srs_s(t), \omega_{\text{amplifying}} fs_b(t)) - ps_b(t)] \Delta t$$

LP3 Feeling b

$$df_{sb}/dt = \eta_{fs_b} [c_{fs_b} (\omega_{\text{feeling}} ps_b) - fs_b]$$

$$fs_b(t + \Delta t) = fs_b(t) + \eta_{fs_b} [c_{fs_b} (\omega_{\text{feeling}} ps_b(t)) - fs_b(t)] \Delta t$$

The state symbols are explained as follows:

ss_s sensor state for stimulus s

srs_s sensory representation state for stimulus s

ps_b preparation state for emotional response b

fs_b feeling state for b

Box 3.1 Numerical representation of a temporal-causal network model for the emotion-feeling cycle

The biological mechanisms briefly sketched above have been used as inspiration for computational mechanisms in earlier work as well, for example, described in Bosse et al. (2008a, b, 2012b). Here in Bosse et al. (2008a, b) a qualitative model without cycles is described; the connection from feeling to preparation was not covered. In Bosse et al. (2012a) the focus is on emotion reading in a social context. A more complex model is presented of which the cycle shown in Fig. 3.1 is part. The computational model specification in both cases (Bosse et al. 2008a, b, 2012b) is not in terms of differential equations as above, but in terms of the hybrid temporal-causal LEADSTO format which covers differential equations as well; e.g., Bosse et al. (2007).

3.3 Emotion Regulation

Controlling or regulating your emotion is often associated to suppressing an emotional response, for example, expressing a neutral poker face. This type of controlling emotions is sometimes considered not very healthy, and a risk for developing serious medical problems. However, it has been found that the mechanisms to regulate emotions form a much wider variety. For example, closing or covering your eyes when a movie is felt as too scary, or avoiding an aggressive person are different forms of control. Emotion regulation mechanisms (e.g., Gross 1998; Goldin et al. 2008) cover *antecedent-focused regulation* (e.g., selection and modification of the situation, attentional deployment, and reappraisal) and *response-focused regulation* (suppression of a response). Examples of antecedent-focused mechanisms are closing your eyes or turning away your gaze from stimuli that trigger too high levels of emotions, redirecting attention, or changing the cognitive interpretation of the situation. Response-focused emotion regulation mechanisms suppress the emotional responses without taking away or modulating the triggers.

Expressing a poker face or fighting against tears are examples of such mechanisms. Emotion regulation mechanisms are processes with a cyclic character. In modeling emotion regulation, in the first place a control state is needed to detect whether an undesired level of emotion occurs. This is assumed to be realised in the prefrontal cortex (Goldin et al. 2008). When this control state has a high activation level (for example, indicating too high levels of an undesired emotion), this can affect a number of other states. Response-focused mechanisms can be modeled by suppressing connections (with negative weight factors) from the control state to preparation and/or effector states. Antecedent-focused mechanisms can be modeled by suppressing connections from the control state to sensor states, sensory representation states or interpretation states. In the conceptual representation of a temporal-causal network model shown in Fig. 3.2 some of these possibilities are depicted.

In conceptual matrix representation the temporal-causal network model shown in Fig. 3.1 is described by Table 3.1. The conceptual representation of the model for emotion regulation as depicted in Fig. 3.2 (or Table 3.2) has been used to obtain a numerical representation of the model according to the systematic approach described in the Chap. 2, as shown in Box 3.2.

Note that here the connection weights from the control state cs_b to other states have a negative value. These connections suppress the activation levels of the destination nodes. The more negative these connection weights are, the stronger the suppression.

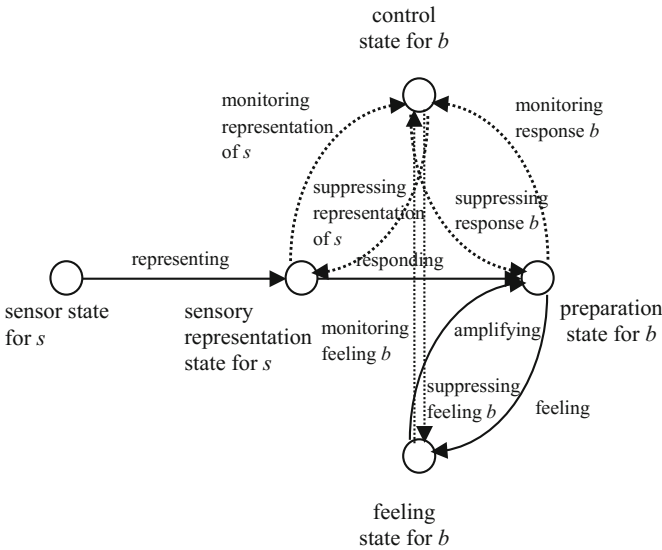


Fig. 3.2 Conceptual representation of a temporal-causal network model showing the cyclic mechanisms for emotion regulation

Table 3.2 Conceptual matrix representation for the model for emotion regulation shown in Fig. 3.2

To From	ss_s	srs_s	ps_b	fs_b	cs_b
ss_s		$\omega_{representing}$			
srs_s			$\omega_{responding}$		$\omega_{monitoring_s}$
ps_b				$\omega_{feeling}$	$\omega_{monitoring_response}$
fs_b			$\omega_{amplifying}$		$\omega_{monitoring_feeling}$
cs_b			$\omega_{suppressing_response}$	$\omega_{suppressing_feeling}$	
η_Y	–	η_{srs_s}	η_{ps_b}	η_{fs_b}	η_{cs_b}
$c_Y(\dots)$	–	$c_{srs_s}(V)$	$c_{ps_b}(V_1, V_2, V_3)$	$c_{fs_b}(V_1, V_2)$	$c_{cs_b}(V_1, V_2, V_3)$

LP1 Generating and regulating representation of stimulus s

$$dsrs_s/dt = \eta_{srs_s} [c_{srs_s}(\omega_{representing} ss_s, \omega_{suppressing_s} cs_b) - srs_s]$$

$$srs_s(t + \Delta t) = srs_s(t) + \eta_{srs_s} [c_{srs_s}(\omega_{representing} ss_s(t), \omega_{suppressing_s} cs_b(t)) - srs_s(t)] \Delta t$$

LP2 Generating and regulating preparation for response b

$$dps_b/dt = \eta_{ps_b} [c_{ps_b}(\omega_{responding} srs_s, \omega_{amplifying} fs_b, \omega_{suppressing_response} cs_b) - ps_b]$$

$$ps_b(t + \Delta t) = ps_b(t)$$

$$+ \eta_{ps_b} [c_{ps_b}(\omega_{responding} srs_s(t), \omega_{amplifying} fs_b(t), \omega_{suppressing_response} cs_b(t)) - ps_b(t)] \Delta t$$

LP3 Generating and regulating feeling b

$$dfs_b/dt = \eta_{fs_b} [c_{fs_b}(\omega_{feeling} ps_b, \omega_{suppressing_feeling} cs_b) - fs_b]$$

$$fs_b(t + \Delta t) = fs_b(t)$$

$$+ \eta_{fs_b} [c_{fs_b}(\omega_{feeling} ps_b(t), \omega_{suppressing_feeling} cs_b(t)) - fs_b(t)] \Delta t$$

LP4 Monitoring stimulus s and response and feeling b

$$dcs_b/dt = \eta_{cs_b} [c_{cs_b}(\omega_{monitoring_s} srs_s, \omega_{monitoring_response} ps_b, \omega_{monitoring_feeling} fs_b) - cs_b]$$

$$cs_b(t + \Delta t) = cs_b(t)$$

$$+ \eta_{cs_b} [c_{cs_b}(\omega_{monitoring_s} srs_s(t), \omega_{monitoring_response} ps_b(t), \omega_{monitoring_feeling} fs_b(t)) - cs_b(t)] \Delta t$$

The state symbols are explained as follows:

ss_s sensor state for stimulus s

srs_s sensory representation state for stimulus s

ps_b preparation state for emotional response b

fs_b feeling state for b

cs_b control state for b

Box 3.2 Numerical representation for the temporal-causal network model for emotion regulation

Biological mechanisms for emotion regulation as discussed also have been the inspiration for computational mechanisms, for example, in Chow et al. (2005), which takes homeostatic principles as a point of departure to address emotion regulation computationally. Also in Bosse et al. (2010a, b, c) such principles are the underlying assumptions, and they are applied to the different phases considered by Gross (1998): situation selection, situation modification, attention deployment, reappraisal, response suppression. Here a different type of specification is used, in LEADSTO format (e.g., Bosse et al. 2007). In Chap. 9 a temporal-causal network model for (reduced) social interaction is presented in which emotion regulation is used for cases of enhanced sensory processing sensitivity to avoid stimuli that are felt as having a too strong impact; also see Treur (2011c, d). This model uses a similar mechanism as described here, but for a specific type of regulation. The same applies to the adaptive temporal-causal network model for dreaming discussed in Chap. 5; here emotion regulation is used to down-regulate fear in dream episodes; see also Treur (2011a). Within Abro et al. (2015) and Manzoor et al. (2016) temporal-causal network models integrating different emotion regulation strategies and for decision making about emotion regulation strategies are discussed. Within Abro et al. (2014) a computational model of the relation between regulation of negative emotions and mood is presented.

3.4 Interaction Between Cognitive and Affective States

Usually it is assumed that behaviour can be described in relation to cognitive states such as beliefs and desires, while leaving affective states aside. The latter types of states are considered as being part of a separate line of (affective) processes that produce their own output, for example, in the sense of emotions and expressions of them. However, this assumed separation between cognitive and affective processes is questioned more and more. Specific examples of questions about such interactions are: how does desiring relate to feeling, and in how far do sensing and believing relate to feeling? Recent neurological findings suggest that this separation of processes may not be a fruitful way to go. For example, as also illustrated in Chap. 1, Sect. 1.2, Phelps (2006) states:

The mechanisms of emotion and cognition appear to be intertwined at all stages of stimulus processing and their distinction can be difficult. (...) Adding the complexity of emotion to the study of cognition can be daunting, but investigations of the neural mechanisms underlying these behaviors can help clarify the structure and mechanisms (Phelps 2006, pp. 46–47).

Similar claims have been made by Pessoa (2008). In experimental contexts different types of effects of affective states on cognitive states have indeed been found; see, for example, Eich et al. (2000), Forgas et al. (2009) and Winkielman et al. (2009). Moreover, in the rapidly developing area of cognitive neuroscience (e.g., Purves et al. 2008; Gazzaniga 2009) more in general knowledge has been contributed on mechanisms for the interaction and intertwining of affective and cognitive states and processes (for example, involving emotion, mood, beliefs or memory); see, for

example, Dolan (2002), LaBar and Cabeza (2006), Pessoa (2008), Phelps (2006) and Storbeck and Clore (2007).

To become a bit more specific, the interaction between beliefs and emotions is discussed in a bit more detail. For example, in Damasio (1999, 2003) it is described how in a person a belief state induces emotions felt within this person:

Even when we somewhat misuse the notion of feeling – as in “I feel I am right about this” or “I feel I cannot agree with you” – we are referring, at least vaguely, to the feeling that accompanies the idea of believing a certain fact or endorsing a certain view. This is because believing and endorsing *cause* a certain emotion to happen. (...) Through either innate design or by learning, we react to most, perhaps all, objects with emotions, however weak, and subsequent feelings, however feeble (Damasio 2003, p. 93).

For the sake of simplicity it is assumed that beliefs are cognitive states representing knowledge about the world and generated (mainly) on the basis of sensing; however, the further processing discussed here does not depend on this. So, for the case of beliefs as cognitive states, during the process that they are generated, beliefs trigger emotional responses that result in certain feelings. However, the process of generation of a cognitive state such as a belief is not fully independent of such associated feelings, as also put forward by Frijda (1993), Lewis (1996), Frijda et al. (2000) and Spinoza (1677):

Beliefs thus are regarded as one of major determinants of emotion, and therefore an important part of the study of emotion can properly be seen as falling under the umbrella of cognitive psychology. Oddly enough, however, the reverse direction of influence in the relation between emotion and cognition has received scant attention. (...) Indeed, such an influence has traditionally been considered to be one of the most important things to be said about emotions. Spinoza (1677/1989) defined emotions as “states that make the mind inclined to think one thing rather than another”. (...) The general proposal thus is that emotions can awaken, intrude into, and shape beliefs, by creating them, by amplifying or altering them, and by making them resistant to change (Frijda et al. 2000, p. 1, 5).

Support for a connection from feeling to belief can be found as well in Damasio’s Somatic Marker Hypothesis; Damasio (1994, 2003), Bechara and Damasio (2005). This is a theory on decision making which provides a central role to emotions felt. Each decision option induces (via an emotional response) a feeling which is used to mark the option. A negative marker has a weakening effect and a positive marker a strengthening effect for the option. Usually the Somatic Marker Hypothesis is applied to provide endorsements or valuations for options for a person’s actions. However, it may be considered plausible that such a mechanism is applicable to valuations of internal states such as beliefs as well. In summary, some indications can be found for the assumption that a belief generates emotional responses and related feelings, and these feelings in turn affect the belief. This provides a pattern based on two cycles in the conceptual representation of a temporal-causal network model depicted in Fig. 3.3. This shows how the cognitive and affective processes are intertwined. Note that the network shown in Fig. 3.3 applies as well to multiple cognitive states active at the same time and multiple emotional responses and feelings. In such a case the level of a given emotion can be affected by the levels of more than one cognitive state by some combination function and similarly the level of a given cognitive state can be affected by the levels of more than one emotion.

Similar analyses can be made for other types of cognitive states. For example, desires are often considered cognitive states with the function of focusing the behaviour by constraining or indicating the options for actions to be chosen. Yet, there is much more to the process of ‘desiring’, especially concerning the feelings associated to it. Desires lead to activations for responses in the form of preparations for certain actions (to fulfill the desire) and their related emotions. Such responses in turn relate in a reciprocal manner to feelings, via cyclic as-if body loops as discussed above. For example, a desire to have some food may trigger a preparation to take some chocolate, which by an as-if body loop in a cyclic manner goes hand in hand with activation of some feeling. This feeling can strengthen both the desire and the preparation. The two cycles shown in Fig. 3.3 model these processes.

A third type of cognitive state considered is a sensory representation. Such a state is closely related to a sensor state and at least this type of state may be believed not to be affected by affective states. However, even here recently findings have been reported suggesting that this independence of affective states cannot be claimed. In particular, in Gazzola et al. (2012) it is reported how for heterosexual men one and the same stimulus (a leg being touched in an invisible manner, by a woman) leads to different sensory activation levels depending on a presented video of either a woman or a man. Such findings suggest that in a diagram as depicted in Fig. 3.3, also an arrow from feeling to sensory representation state can be drawn (Table 3.3).

A numerical representation for this temporal-causal network model for the interaction between (activation levels of) cognitive and affective states can be found in Box 3.3. Here the connection weights are assumed nonnegative. Note that the higher the values $\omega_{\text{interpreting}}$, $\omega_{\text{emotion_integration}}$ and $\omega_{\text{responding}}$ of the weights of the connections from and to the cognitive state c , and the values $\omega_{\text{amplification}}$ and ω_{feeling} for the connections between the emotional response and feeling state, the higher the activation value of the cognitive state c .

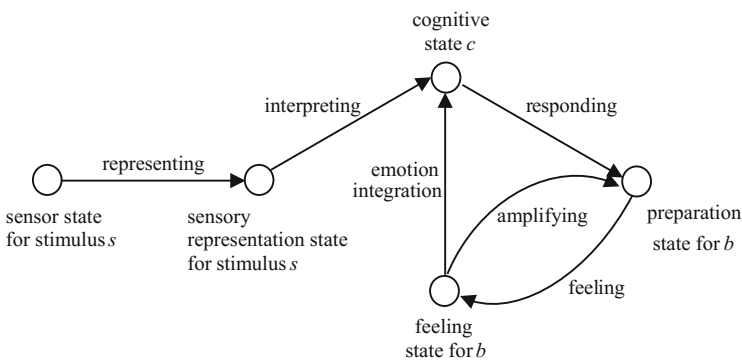


Fig. 3.3 Graphical conceptual representation of a temporal-causal network model of the cyclic processes of mutual interaction between cognitive and affective states

Table 3.3 Conceptual matrix representation of the temporal-causal network model for interaction between cognitive and affective states

To From	ss_s	srs_s	c	ps_b	fs_b
ss_s		$\omega_{\text{representing}}$			
srs_s			$\omega_{\text{interpreting}}$		
c				$\omega_{\text{responding}}$	
ps_b					ω_{feeling}
fs_b			$\omega_{\text{emotion_integration}}$	$\omega_{\text{amplifying}}$	
η_Y	–	η_{srs_s}	η_c	η_{ps_b}	η_{fs_b}
$c_Y(\dots)$	–	$c_{srs_s}(V)$	$c_c(V_1, V_2)$	$c_{ps_b}(V_1, V_2)$	$c_{fs_b}(V)$

LP1 Representing stimulus s

$$dsrs_s/dt = \eta_{srs_s} [c_{srs_s}(\omega_{\text{representing}}ss_s) - srs_s]$$

$$srs_s(t + \Delta t) = srs_s(t) + \eta_{srs_s} [c_{srs_s}(\omega_{\text{representing}}ss_s(t)) - srs_s(t)]\Delta t$$

LP2 Generating interpretation c

$$dc/dt = \eta_c [c_c(\omega_{\text{interpreting}}srs_s, \omega_{\text{emotion_integration}}fs_b) - c]$$

$$c(t + \Delta t) = c(t) + \eta_c [c_c(\omega_{\text{interpreting}}srs_s(t), \omega_{\text{emotion_integration}}fs_b(t)) - c(t)]\Delta t$$

LP3 Generating and amplifying response b

$$dps_b/dt = \eta_{ps_b} [c_{ps_b}(\omega_{\text{responding}}c, \omega_{\text{amplifying}}fs_b) - ps_b]$$

$$ps_b(t + \Delta t) = ps_b(t) + \eta_{ps_b} [c_{ps_b}(\omega_{\text{responding}}c(t), \omega_{\text{amplifying}}fs_b(t)) - ps_b(t)]\Delta t$$

LP4 Feeling b

$$dfs_b/dt = \eta_{fs_b} [c_{fs_b}(\omega_{\text{feeling}}ps_b) - fs_b]$$

$$fs_b(t + \Delta t) = fs_b(t) + \eta_{fs_b} [c_{fs_b}(\omega_{\text{feeling}}ps_b(t)) - fs_b(t)]\Delta t$$

The symbols are explained as follows:

ss_s sensor state for stimulus s

srs_s sensory representation state for stimulus s

ps_b preparation state for emotional response b

fs_b feeling state for b

c cognitive interpretation state c

Box 3.3 Numerical representation for a temporal-causal network model for interaction between cognitive and affective states

The interaction between belief and feeling has also been addressed in Memon and Treur (2010), using a temporal-causal network model specified in LEADSTO format (Bosse et al. 2007). Moreover, in Memon and Treur (2010) the connection to the belief is adaptive. The type of combination function $c_{ps_b}(\dots)$ used in Memon and Treur (2010) is:

$$\begin{aligned} c_{ps_b}(V_1, V_2) &= \mathbf{aproduct}_\beta(V_1, V_2) \\ &= \beta \mathbf{cproduct}(V_1, V_2) + (1 - \beta) \mathbf{product}(V_1, V_2) \\ &= \beta(1 - (1 - V_1)(1 - V_2)) + (1 - \beta)V_1V_2 \end{aligned}$$

with β a bias parameter in $[0, 1]$. The interaction between desire and feeling has been worked out in more detail in an adaptive temporal-causal network model in Bosse et al. (2010a, b, c). Here adaptivity based on feedback of actual execution is part of the model. Also this model was specified in LEADSTO format.

3.5 Emotion-Related Valuing in Decision-Making

In the area of decision making the role of emotions has been discussed since long. From an idealised rationality perspective it has long been assumed that emotions can only disturb proper rational decision making and should be left out of the process in order to come up with adequate decisions. However, this has been questioned in more recent times. For example, in Loewenstein and Lerner (2003, p. 619) it is claimed that recent research points at the positive functions served by emotions; see also the quote in Chap. 1, Sect. 1.2. Can you make an adequate decision without feeling good about it? If you make a decision with a bad feeling this may cast doubt on how robust the decision is: at any occasion in the (near) future you may be tempted to change it into a different decision. The area of decision making is another specific area in which affective elements play an important role. The focus in decision making is on how to perform valuing of situations or options for actions to be decided for. More specifically, feelings generated in relation to an observed situation and prepared action option play an important role in valuing predicted or imagined effects of such an action in the situation. Such valuations have been related to amygdala activations (see, e.g., Morrison and Salzman 2010; Murray 2007; Salzman and Fusi 2010). Although traditionally an important function attributed to the amygdala concerns the context of fear, in recent years much evidence on the amygdala in humans has been collected showing a function beyond this fear context.

A model for such a decision process can be designed based on the following cyclic processes. A preparation state for one or more actions a is triggered by a sensed stimulus s , and by internal simulation a prediction is made of the effect e of such an action a :

sensor state for stimulus $s \rightarrow$ sensory representation of $s \rightarrow$ preparation state for $a \rightarrow$
 sensory representation of predicted effect e

A preparation for an emotional response b is triggered by this predicted effect e , and via a cyclic as-if body loop a feeling for this emotion is generated:

sensory representation of $e \rightarrow$ preparation state for emotion $b \rightarrow$ feeling state for b
 feeling state for $b \rightarrow$ preparation state for emotion b

Feeling this emotion represents a way of experiencing the value of the predicted effect e of action a : to which extent it is felt as positive. This valuation in turn affects the activation of the concerning option in the sense that the more positive this feeling, the more is the preparation for a strengthened:

feeling state for $b \rightarrow$ preparation state for a

This adds an extra cycle in the process, so this pattern involves two cycles. See also Chap. 6 for more details (e.g., Sect. 6.2, Fig. 3.1). In that chapter also a simplified but adaptive model is discussed, based on Treur and Uthair (2011, 2015). In particular, it is analysed how adaptivity can be added and in how far this makes the model behave rationally for a given environment.

3.6 Emotions and Social Contagion

Emotions also play an important role in mutual social interactions. In a social context usually emotions of different individuals affect each other: emotion contagion. The question that may arise, for example, is in how far an individual in a group free is in having his or her own emotions. Moreover, the role of emotion contagion may be considered for the miracle that groups with individuals with initially different preferences often come to coherent common decisions and all members feel good with these decisions.

The mechanisms underlying emotion contagion can be considered in a more detailed manner. From the area of Social Neuroscience it has been found that mirror neurons and internal simulation are key elements in these mechanisms. Mirror neurons are neurons that do not only have the function to prepare for a certain action or body change (e.g., a face expression), but are also activated upon observing somebody else who is performing this action or body change. They have been found both in monkeys and humans; e.g., Rizzolatti and Sinigaglia (2008), Iacoboni (2008) and Mukamel et al. (2010). Mirror neurons make that some specific sensory input (an observed person) directly links to activation of related preparation states.

Viewed from a distance, an emotion contagion cycle for two persons A and B can be described by causal chains as follows. Here an emotion b shown is not only affected by a stimulus s from the (non-human) environment, but also by sensing emotions shown by other individuals. First, person A responds by showing

emotion b (for example, a smile) both to a stimulus s (for example, a funny picture) or to observing person B showing b :

sensor state of person A for stimulus $s \rightarrow$ person A shows b
 sensor state of person A for person B showing $b \rightarrow$ person A shows b

The same applies to person B with respect to person A :

sensor state of person B for stimulus $s \rightarrow$ person B shows b
 sensor state of person B for person A showing $b \rightarrow$ person B shows b

This makes a cyclic process through the two persons, and if more persons are taken into account a combination of multiple cyclic processes through them.

The internal mechanisms behind such social contagion by mirroring based on mirror states and internal simulation are addressed in more detail in Chap. 7. In Chap. 11 more detailed adaptive temporal-causal network models for social interaction are addressed, thereby abstracting from internal processes. In Manzoor and Treur (2013, 2015) an integrated temporal-causal network model for emotion regulation and emotion contagion is discussed.

3.7 Discussion

In this chapter a unifying neurologically inspired perspective on the dynamics and interaction of emotions was discussed, making use of knowledge of mechanisms from Cognitive, Affective and Social Neuroscience. The contents of this chapter are mainly based on a number of earlier papers. The part on emotions and feelings in Sect. 3.2 is based on Bosse et al. (2008b, 2010a, b). The part on emotion regulation in Sect. 3.3 is based on Treur (2011a, b, 2014), Bosse et al. (2013). The part on interaction between cognitive and affective states in Sect. 3.4 is based on Memon and Treur (2010), Bosse et al. (2010a, b). The part on rationality and emotion is based on Treur and Umair (2011). The part on emotion contagion in Sect. 3.5 is based on Bosse et al. (2015).

It was discussed how many cyclic connections in the brain can be found and play an important role in brain processes (see also, Bell 1999; Crick and Koch 1998; Potter 2007; Sporns et al. 2000), and in particular how affective states can have bidirectional associations to many other types of mental states and behaviour (e.g., Critchley 2005; Damasio 2003; Frijda et al. 2000; Scherer 2009). It was indicated how according to a Network-Oriented Modeling approach the type of processes considered can be described by graphical conceptual representations of temporal-causal networks, or by conceptual matrix representations, and by numerical representations in difference and differential equation format, in line with what was discussed in Chap. 2. Thus a unifying integrative dynamical perspective on modeling emotions was obtained that can be used to model how emotions relate to a variety of other mental states and processes such as feelings, beliefs, desires, experiences, and valuations in decision making. Moreover, the approach covers

how emotions of different persons affect each other (emotion contagion). This social element will be discussed in more detail in Chap. 7.

The types of examples of cycles discussed here can be and actually have been integrated further. More advanced temporal-causal network models for emotions and their role in mental functioning may involve different types of cycles that have to be integrated. Examples of this further integration are integration of an emotion regulation cycle with a social interaction cycle as described in Aziz et al. (2011), or the integration of cognitive-affective cycles with social interaction cycles as described in Bosse et al. (2012a) and Hoogendoorn et al. (2011). Furthermore, in a number of models mechanisms for adaptivity and emotional response-feeling cycles have been integrated with other types of cycles (e.g., Bosse et al. 2010a, b, c; Memon and Treur 2010).

Note that the perspective as described here assumes that cognitive and affective states can be distinguished, while their interaction makes the cognitive and affective processes intertwined. However, as discussed, for example, in Samsonovich (2012) as an alternative it is also possible to consider mental states in general that are not distinguishable as either cognitive or affective, but which may have both a cognitive and an affective aspect. This makes an alternative perspective possible where the distinction cognitive/affective can almost disappear.

In subsequent chapters the integration of emotions in different types of mental processes will be addressed in some more detail, in particular for emotions and dreaming (Chap. 4), emotions, dreaming and fear extinction learning (Chap. 5), emotions and rationality in action selection (Chap. 6), and emotions within social processes (Chap. 7).

References

- A.H. Abro, M.C.A. Klein, A.R. Manzoor, S.A. Tabatabaei, J. Treur, A computational model of the relation between regulation of negative emotions and mood, in *Proceedings of the 21th International Conference on Neural Information Processing, ICONIP'14*, eds. by C.K. Loo, Y. Keem Siah, K.K.W. Wong, A.T. Beng Jin, K. Huang. Lecture Notes in Artificial Intelligence, Springer, Lecture Notes in Computer Science, vol. 8834 (2014), pp. 59–68
- A.H. Abro, A.R. Manzoor, S.A. Tabatabaei, J. Treur, A computational cognitive model integrating different emotion regulation strategies, in *Proceedings of the Sixth International Conference on Biologically Inspired Cognitive Architectures, BICA'15*. Procedia Computer Science, vol. 71 (Elsevier Publishers, 2015), pp. 157–168
- A.A. Aziz, J. Treur, C.N. van der Wal, An agent-based model for integrated contagion and regulation of negative mood. in *Agents in Principle, Agents in Practice, Proceedings of the 14th International Conference on Principles and Practice of Multi-Agent Systems, PRIMA'11*, eds. by D. Kinny et al. (Springer, 2011) Lecture Notes in Artificial Intelligence, vol. 7047, pp. 83–96
- W. Becker, A.F. Fuchs, Prediction in the oculomotor system: smooth pursuit during transient disappearance of a visual target. *Exp. Brain Res.* **57**, 562–575 (1985)
- A. Bechara, A. Damasio, The somatic marker hypothesis: a neural theory of economic decision. *Games Econ. Behav.* **52**, 336–372 (2005)

- A. Bell, Levels and loops: the future of artificial intelligence and neuroscience. *Phil. Trans. R. Soc. Lond. B* **354**, 2013–2020 (1999)
- T. Bosse, C.M. Jonker, L. van der Meij, J. Treur, A language and environment for analysis of dynamics by simulation. *Int. J. Artif. Intell. Tools* **16**, 435–464 (2007)
- T. Bosse, C.M. Jonker, J. Treur, Formalisation of Damasio's theory of emotion, feeling and core consciousness. *Conscious. Cogn.* **17**, 94–113 (2008a)
- T. Bosse, Z.A. Memon, J. Treur, Adaptive estimation of emotion generation for an ambient agent model. in *Ambient Intelligence, Proceedings of the Second European Conference on Ambient Intelligence, Aml'08*, E. Aarts, J.L. Crowley, B. de Ruyter, H. Gerhauser, A. Pflaum, J. Schmidt, R. Wichert. *Lecture Notes in Computer Science*, vol. 5355 (Springer, 2008b), pp. 141–156
- T. Bosse, M. Hoogendoorn, Z.A. Memon, J. Treur, M. Umair, An adaptive model for dynamics of desiring and feeling based on Hebbian learning. in *Proceedings of the Second International Conference on Brain Informatics, BI'10*, eds. by Y. Yao, R. Sun, T. Poggio, J. Liu, N. Zhong, J. Huang, *Lecture Notes in Artificial Intelligence*, vol. 6334 (Springer, 2010a), pp. 14–28
- T. Bosse, M. Hoogendoorn, Z.A. Memon, J. Treur, M. Umair, An adaptive model for dynamics of desiring and feeling based on Hebbian learning. in *Proceedings of the Second International Conference on Brain Informatics, BI'10*, eds. by Y. Yao, R. Sun, T. Poggio, J. Liu, N. Zhong, J. Huang, *Lecture Notes in Artificial Intelligence*, vol. 6334 (Springer, 2010b), pp. 14–28. Extended version in *Cognitive Systems Research*, 2012, in press
- T. Bosse, M. Pontier, J. Treur, A computational model based on gross' emotion regulation theory. *Cogn. Syst. Res.* **11**, 211–230 (2010c)
- T. Bosse, M. Hoogendoorn, Z.A. Memon, J. Treur, M. Umair, A computational model for dynamics of desiring and feeling. *Cogn. Syst. Res. J.* **19**, 39–61 (2012a)
- T. Bosse, Z.A. Memon, J. Treur, A cognitive and neural model for adaptive emotion reading by mirroring preparation states and Hebbian learning. *Cogn. Syst. Res.* **12**, 39–58 (2012b)
- T. Bosse, C.G. Gerritsen, J. de Man, J. Treur, Learning emotion regulation strategies: a cognitive agent model, in *Proceedings of the 13th International Conference on Intelligent Agent Technology, IAT'13*, vol. 2 (IEEE Computer Society Press, 2013), pp. 245–252
- T. Bosse, R. Duell, Z.A. Memon, J. Treur, C.N. van der Wal, Agent-based modelling of emotion contagion in groups. *Cogn. Comput. J.* **7**, 111–136 (2015)
- S.-M. Chow, N. Ram, S.M. Boker, F. Fujita, G. Clore, Emotion as a thermostat: representing emotion regulation using a damped oscillator model. *Emotion* **5**, 208–225 (2005)
- F. Crick, C. Koch, Constraints on cortical and thalamic projections: the no-strong-loops hypothesis. *Nature* **391**, 245–250 (1998)
- H.D. Critchley, Neural mechanisms of autonomic, affective, and cognitive integration. *J. Comp. Neurol.* **493**, 154–166 (2005)
- A.R. Damasio, *Descartes' Error: Emotion, Reason and the Human Brain* (Papermac, London, 1994)
- A.R. Damasio, *The Feeling of What Happens. Body and Emotion in the Making of Consciousness* (Harcourt Brace, New York, 1999)
- A.R. Damasio, *Looking for Spinoza* (Vintage books, London, 2003)
- A.R. Damasio, *Self Comes to Mind: Constructing the Conscious Brain* (Pantheon Books, New York, 2010)
- R.J. Dolan, Emotion, cognition, and behavior. *Science* **298**, 1191–1194 (2002)
- E. Eich, J.F. Kihlstrom, G.H. Bower, J.P. Forgas, P.M. Niedenthal, *Cognition and Emotion* (Oxford University Press, New York, 2000)
- E.A. Feigenbaum, Artificial intelligence: themes in the second decade. *Inf. Process.* **68**, 1008–1024 (1969)
- J.P. Forgas, L. Goldenberg, C. Unkelbach, Can bad weather improve your memory? An unobtrusive field study of natural mood effects on real-life memory. *J. Exp. Soc. Psychol.* **45**, 254–257 (2009)
- N.H. Frijda, The place of appraisal in emotion. *Cogn. Emot.* **7**, 357–387 (1993)
- N.H. Frijda, A.S.R. Manstead, S. Bem, The influence of emotions on beliefs, in *Emotions and Beliefs: How Feelings Influence Thoughts*, eds. by N.H. Frijda et al. (Cambridge University Press, 2000), pp. 1–9

- M.S. Gazzaniga, (ed.), *The Cognitive Neurosciences*, 4th edn. (MIT Press, 2009)
- V. Gazzola, M.L. Spezio, J.A. Etzela, F. Castelli, R. Adolphs, C. Keysers, Primary somatosensory cortex discriminates affective significance in social touch. *PNAS* (2012). doi:[10.1073/pnas.1113211109](https://doi.org/10.1073/pnas.1113211109)
- P.R. Goldin, K. McRae, W. Ramel, J.J. Gross, The neural bases of emotion regulation: reappraisal and suppression of negative emotion. *Biol. Psychiatry* **63**, 577–586 (2008)
- A.I. Goldman, *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading* (Oxford University Press, New York, 2006)
- J.J. Gross, Antecedent- and response-focused emotion regulation: divergent consequences for experience, expression, and physiology. *J. Pers. Soc. Psychol.* **74**, 224–237 (1998)
- E. Hatfield, R.L. Rapson, Y.L. Le, Emotional contagion and empathy, in *The Social Neuroscience of Empathy*, ed. by J. Decety, W. Ickes (MIT, Cambridge, 2009)
- G. Hesslow, Conscious thought as simulation of behaviour and perception. *Trends Cogn. Sci.* **6**, 242–247 (2002)
- M. Hoogendoorn, J. Treur, C.N. van der Wal, A. van Wissen, Agent-based modelling of the emergence of collective states based on contagion of individual states in groups. *Trans. Comput. Collect. Intell.* **3**, 152–179 (2011)
- M. Iacoboni, *Mirroring People: The New Science of How We Connect with Others* (Farrar, Straus & Giroux, 2008)
- W. James, What is an emotion. *Mind* **9**, 188–205 (1884)
- K.S. LaBar, R. Cabeza, Cognitive neuroscience of emotional memory. *Nat. Rev. Neurosci.* **7**, 54–64 (2006)
- R. Lazarus, Progress on a cognitive-motivational-relational theory of emotion. *Am. Psychol.* **46**, 819–834 (1991)
- M.D. Lewis, Self-organising cognitive appraisals. *Cogn. Emot.* **10**, 1–25 (1996)
- G.F. Loewenstein, J.S. Lerner, The role of affect in decision making, in *Handbook of Affective Sciences*, ed. by R.J. Davidson, K.R. Scherer, H.H. Goldsmith (Oxford University Press, Oxford, 2003), pp. 619–642
- A.R. Manzoor, J. Treur, Modelling the role of emotion regulation and contagion in socially affected decision making, in *Proceedings of the 9th International Conference on Cognitive Science, ICCS'13*, eds. by C.S. Teh, H.R. Chae, S.A.Z. Adruce, P.N. Anding, C.J. Chen, N.A. Aziz, K.W. Tan. *Procedia Social and Behavioral Sciences*, vol. 97 (Elsevier, 2013), pp. 73–82
- A. Manzoor, J. Treur, An agent-based model for integrated emotion regulation and contagion in socially affected decision making. *Biol. Inspir. Cogn. Archit. J.* **12**(2015), 105–120 (2015)
- A. Manzoor, A.H. Abro, J. Treur, Monitoring the Impact of Negative Events and Deciding about Emotion Regulation Strategies (2016)
- Z.A. Memon, J. Treur, On the reciprocal interaction between believing and feeling: an adaptive agent modelling perspective. *Cogn. Neurodyn.* **4**, 377–394 (2010)
- S.E. Morrison, C.D. Salzman, Re-valuing the amygdala. *Curr. Opin. Neurobiol.* **20**, 221–230 (2010)
- R. Mukamel, A.D. Ekstrom, J. Kaplan, M. Iacoboni, I. Fried, Single-neuron responses in humans during execution and observation of actions. *Curr. Biol.* **20**, 750–756 (2010)
- E.A. Murray, The amygdala, reward and emotion. *Trends Cogn. Sci.* **11**, 489–497 (2007)
- U. Neisser, The Imitation of man by machine. *Science* **139**, 193–197 (1963)
- C. Peck, J. Kozloski, The computational basis of emotions and implications for cognitive architectures, in *Proceedings of the Second International Conference on Biologically Inspired Cognitive Architectures, BICA'11*, eds. by A.V. Samsonovich, K.R. Johansson. *Frontiers in Artificial Intelligence and Applications*, vol. 233 (IOS Press, 2011), pp. 269–281
- L. Pessoa, On the relationship between emotion and cognition. *Nat. Rev. Neurosci.* **9**, 148–158 (2008)
- E.A. Phelps, Emotion and cognition: insights from studies of the human Amygdala. *Annu. Rev. Psychol.* **57**, 27–53 (2006)
- S.M. Potter, What can artificial intelligence get from neuroscience? in *Artificial Intelligence Festschrift: The Next 50 Years*, eds. by M. Lungarella, J. Bongard, R. Pfeifer (Springer, Berlin, 2007)

- D. Purves, E.M. Brannon, R. Cabeza, S.A. Huettel, K.S. LaBar, M.L. Platt, M.G. Woldorff, *Principles of Cognitive Neuroscience* (Sinauer Associates Inc., Sunderland, 2008), p. 2008
- G. Rizzolatti, C. Sinigaglia, *Mirrors in the Brain: How Our Minds Share Actions and Emotions* (Oxford University Press, 2008)
- I.J. Roseman, Appraisal determinants of emotions: constructing a more accurate and comprehensive theory. *Cogn. Emot.* **10**, 241–278 (1996)
- C.D. Salzman, S. Fusi, Emotion, cognition, and mental state representation in amygdala and prefrontal cortex. *Annu. Rev. Neurosci.* **33**, 173–202 (2010)
- A.V. Samsonovich, An approach to building emotional intelligence in artifacts, in *Proceedings of the 8th International Conference on Cognitive Robotics (CogRob'12): Papers from the 2012 AAAI Workshop*, eds. by M. Pagnucco, W. Burgard, K. Konolige, S. Vassos, AAAI Technical Report WS-12-03 (AAAI Press, Menlo Park, 2012)
- K.R. Scherer, On the sequential nature of appraisal processes: indirect evidence from a recognition task. *Cogn. Emot.* **13**, 763–793 (1999)
- K.R. Scherer, Emotions are emergent processes: they require a dynamic computational architecture. *Phil. Trans. R. Soc. B* **364**, 3459–3474 (2009)
- H.A. Simon, Motivational and emotional controls of cognition. *Psychol. Rev.* **74**, 29–39 (1967)
- B. Spinoza, *Ethica*. (translated by G.H.R. Parkinson) (Everyman, London, 1677/1989)
- O. Sporns, G. Tononi, G.M. Edelman, Connectivity and complexity: the relationship between neuroanatomy and brain dynamics. *Neural Networks* **13**, 909–922 (2000)
- J. Storbeck, G.L. Clore, On the interdependence of cognition and emotion. *Cogn. Emot.* **21**, 1212–1237 (2007)
- J. Treur, A computational agent model using internal simulation to generate emotional dream episodes, in *Proceedings of the Second International Conference on Biologically Inspired Cognitive Architectures, BICA'11*, eds. by A.V. Samsonovich, K.R. Jóhannsdóttir. *Frontiers in Artificial Intelligence and Applications*, vol. 233 (IOS Press, 2011a), pp. 389–399
- J. Treur, Dreaming your fear away: a computational model for fear extinction learning during dreaming, in *Proceedings of the 18th International Conference on Neural Information Processing, ICONIP'11, Part III*, eds. by B.-L. Lu, L. Zhang, J. Kwok, *Lecture Notes in Artificial Intelligence*, vol. 7064 (Springer, Berlin, 2011b), pp. 197–209
- J. Treur, A cognitive agent model displaying and regulating different social response patterns. in *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence, IJCAI'11*, ed. by T. Walsh (2011c), pp. 1735–1742
- J. Treur, From mirroring to the emergence of shared understanding and collective power (invited talk). in *Proceedings of the hird International Conference on Computational Collective Intelligence, ICCCI'11, Part I*, eds. by P. Jedrzejowicz, N.T. Nguyen, K. Hoang. *Lecture Notes in Artificial Intelligence*, vol. 6922 (Springer, 2011d), pp. 1–16. Extended version in *Transactions on Computational Collective Intelligence*, vol. 8 (2012), pp. 168–191
- J. Treur, Displaying and regulating different social response patterns: a computational agent model. *Cogn. Comput. J.* **6**(2014), 182–199 (2014)
- J. Treur, M. Umair, On rationality of decision models incorporating emotion-related valuing and hebbian learning, in *Proceedings of the 18th International Conference on Neural Information Processing, ICONIP'11, Part III*, eds. by B.-L. Lu, L. Zhang, J. Kwok. *Lecture Notes in Artificial Intelligence*, vol. 7064 (Springer, 2011), pp. 217–229
- J. Treur, M. Umair, Emotions as a vehicle for rationality: rational decision making models based on emotion-related valuing and Hebbian learning. *Biol. Inspir. Cogn. Arch. J.* **14**(2015), 40–56 (2015)
- P. Winkelman, P.M. Niedenthal, L.M. Oberman, Embodied perspective on emotion-cognition interactions, in *Mirror Neuron Systems: The Role of Mirroring Processes in Social Cognition*, ed. by J.A. Pineda (Humana Press/Springer Science, 2009), pp. 235–257

Chapter 4

How Do You Feel Dreaming

Using Internal Simulation to Generate Emotional Dream Episodes

Abstract In this chapter a Network-Oriented Modeling approach is applied to model dreaming based on internal simulation. Building blocks for this internal simulation are memory elements in the form of sensory representations and their associated emotions. In the presented temporal-causal network model, under influence of associated feeling levels and mutual competition, some sensory representation states pop up in different dream episodes. The activation levels of both the feeling and the sensory representation states are regulated by control states. The model was evaluated by a number of simulation experiments for different scenarios.

4.1 Introduction

The mechanisms and functions of dreaming have received much attention in the recent cognitive and neurological literature; e.g., Hobson (2009), Levin and Nielsen (2007, 2009), Revonsuo (2000), Valli et al. (2005); Valli and Revonsuo (2009), Windt and Noreika (2011), Yoo et al. (2007), Kahn et al. (2013), Deliens et al. (2014), Goldstein and Walker (2014). As often negative emotions play an important role in dreams, this aspect is also addressed in some depth, especially in the context of improving skills for coping with threatening situations (e.g., Revonsuo 2000; Valli et al. 2005; Valli and Revonsuo 2009) or strengthening regulation of fear emotions by what is called fear extinction learning (e.g., Levin and Nielsen 2007; Walker and van der Helm 2009; van der Helm et al. 2011; Rosales-Lagarde et al. 2012; Markarian et al. 2013; Mauss et al. 2013; Pace-Schott et al. 2015). Abstracting from more specific context or purpose, a more general perspective present in dream literature as mentioned, is that dreaming can be considered a form of internal simulation of real-life-like processes as a form of training in order to learn, adapt or improve capabilities, which would be less easy to achieve in real life.

In this chapter a Network-Oriented Modeling approach based on temporal-causal network models is presented that addresses the type of internal simulation that is assumed to take place in dreaming. For the different episodes, the internal simulation incorporates interrelated processes of activation of sensory representation states

(from memory) providing mental images, and activation of associated feelings. Moreover, the model uses a mechanism for emotion regulation to suppress the feeling levels and the sensory representation states.

The structure of the chapter is as follows. In Sect. 4.2 the basic concepts used are briefly introduced. In Sect. 4.3 the temporal-causal network model is described in more detail. Section 4.4 discusses simulation results providing dream scenarios. In Sect. 4.5 the relation of the model with neurological theories and findings is addressed. Finally, Sect. 4.6 is a discussion.

4.2 Memory Elements, Emotions and Internal Simulation in Dreaming

In this section it is discussed how in dreaming memory elements with their associated emotions are used as building blocks for an internal simulation of real life.

Using memory elements and their emotional associations Within the literature the role of memory elements providing content for dreams is well-recognized; e.g.:

... dreaming tends to express memory elements as though original memories had been reduced to more basic units (...). Often, these appear as isolated features, such as an attribute of a familiar place or character (e.g., “there was a stranger who had my mother’s style of hair”) (Levin and Nielsen 2007, p. 499).

The role of emotional aspects in activating such memory elements is emphasized; e.g.:

...elements may be activated as a function of emotional concerns (...) but with the possible introduction of some pseudorandom and incompatible associations (Levin and Nielsen 2007, p. 500).

In particular, it is recognized that the choice for memory elements with some emotional association and (re)combining them into a dream facilitates fear generation:

During dreaming, conjunctive representations are rendered into virtual simulations or “here-and-now” illusions (Nielsen and Stenstrom 2005) to maximize their impact upon the amygdala, which tends to respond to perceptual, rather than imaginal, stimuli (Levin and Nielsen 2007, p. 500).

The emotional associations of the sensory memory elements may make that a person has to cope with high levels of emotions (e.g., fear) felt in the dream. *Emotion regulation* mechanisms are used to control emotions that are felt as too strong; e.g., Goldin et al. (2008), Gross (1998, 2007). Such mechanisms cover *antecedent-focused regulation* (e.g., selection and modification of the situation, attentional deployment, and reappraisal) and *response-focused regulation* (suppression of a response).

Dreaming as internal simulation Dreams can be considered as flows of activated sequences of images based on (re)combined memory elements:

Recombinations of memory elements give dreams at once their alien and their familiar quality. (...) the new image sequences consist, for the most part, of lifelike simulations of first-person reality. Memory elements are recombined (..) to produce coherent, continuous simulations of waking life experience (Levin and Nielsen 2007, p. 500).

Such flows can be related to the notion of *internal simulation* put forward, among others, by Hesslow (1994, 2002, 2012), Damasio (1994, 1999), Goldman (2006), Barsalou (2009), Marques and Holland (2009), Pezzulo et al. (2013). The idea of internal simulation is that sensory representation states are activated (e.g., mental images), which in response trigger associated preparation states for actions or bodily changes, which, by prediction links, in turn activate other sensory representation states.

sensory representation states → preparation states → sensory representation states

The latter states represent the effects of the prepared actions or bodily changes, without actually having executed them. Being inherently cyclic, the simulation process can go on indefinitely. Internal simulation has been used, for example, to describe (imagined) processes in the external world, e.g., prediction of effects of own actions (Becker and Fuchs 1985), or processes in another person's mind, e.g., emotion recognition or mindreading (Goldman 2006) or processes in a person's own body (Damasio 1994). Although usually internal simulation as briefly described above concerns mental processes for awake persons, it is easy to imagine that it may be applicable as well to describe dreaming.

Feeling emotions by internal simulation of body states The idea of internal simulation has been exploited in particular by applying it to bodily changes expressing emotions, using the notion of *as-if body loop* (Damasio 1994). For more details on the role of body loops and as-if body loops in emotions, see Chap. 3.

On purposes of dreaming as internal simulation One theory explicitly referring to a purpose of dreaming as internal simulation is the threat simulation theory of the evolutionary function of dreaming (e.g. Revonsuo 2000; Valli et al. 2005; Valli and Revonsuo 2009). This theory assumes that dreaming is an evolutionary adaptation to be able to rehearse coping with threatening situations in a safe manner. Others consider the function of dreaming in strengthening the emotion regulation capabilities for fear; e.g., Levin and Nielsen (2007, 2009), Franzen et al. (2009), Gujar et al. (2011), Walker (2009), Walker and van der Helm (2009), Yoo et al. (2007), van der Helm et al. (2011), Rosales-Lagarde et al. (2012), Mauss et al. (2013), Goldstein and Walker (2014), Pace-Schott et al. (2015). For this perspective, the purpose of dreaming is to improve the coping with the own fear emotions in real life. For both purposes adequate exercising material is needed for the dreams: fearful situations have to be imagined, built on memory elements suitable for fear arousal. The temporal-causal network model presented in Sect. 4.3 provides this, but it abstracts from the purpose; it does not commit to any of the purposes mentioned.

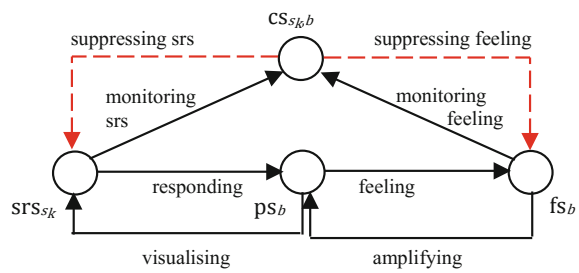
4.3 A Temporal-Causal Network Model Generating Dream Episodes

The temporal-causal network model presented here models the mechanisms discussed in Sect. 4.2. It is meant to address scenarios of the following type:

- *Fearful stimulus*
A (maybe traumatic) stimulus s_1 is given for which previously a high extent of fear has been developed, and for which from time to time a sensory representation state is triggered by memory (for the model this is considered an external trigger)
- *Emotional response*
The activation of the sensory representation of s_1 leads to preparation for a bodily fear response b , and by an as-if body loop to an enhanced feeling level based on b
- *Emotion regulation*
By emotion regulation the sensory representation of s_1 and the feeling state are suppressed: both the experience of fear, and the activation level of the sensory representation of s_1 become low; also no episode state for s_1 occurs, as this is blocked due to the traumatic event
- *Visualisation of the fear*
Other fear-associated stimuli s_k for $k \geq 2$ are available for which the person has less strong previous experiences; the sensory representation states for these s_k are activated by links from the preparation state for b , depending on the strength of these links; this can be viewed as visualisation of the fear
- *Emotional response to visualisation*
When the sensory representation state of a stimulus s_k is activated, this leads to an enhanced activation level of the preparation state for b
- *Experiencing more fear*
Due to the higher activation level of preparation for b , via the as-if body loop also the feeling level for b becomes higher: the person experiences more fear
- *Stronger emotion regulation*
By the control states for emotion regulation for an active sensory representation for s_k both the fear feeling level and the sensory activation level of s_k are suppressed
- *Competition for dream episodes*
The active sensory representations for s_k lead to corresponding dream episode states, which are in competition with each other by mutual inhibition to get dominance as a dream episode

In Fig. 4.1 the basic model for a given sensory representation state srs_{s_k} is shown. It shows emotion generation via emotional response preparation state ps_b and feeling state fs_b (as-if body loop) and emotion regulation through control state $\text{cs}_{s_k,b}$ suppressing the feeling state fs_b and the given sensory representation state srs_{s_k} ; a summary of the states used is shown in Table 4.1. The inhibiting links are

Fig. 4.1 Graphical conceptual representation of a temporal-causal network model for generation and regulation



indicated by dotted arrows (in red). The two links between srs_{s_k} and ps_b indicate the association between stimulus s_k and emotional response b ; the link from ps_b to srs_{s_k} indicates the (predictive link) which is a basis for the internal simulation: e.g., the emotion triggers a certain (expected) mental image. The links between ps_b and fs_b indicate an as-if body loop.

As shown in Table 4.1 a dream episode state for s_k is indicated by des_{s_k} , moreover the trigger for srs_{s_k} from memory is indicated by mt_{s_k} ; this will be applied for s_1 . Note that in Fig. 4.1 a sensory representation state for only one stimulus s_k is depicted. In the specification of the model below an arbitrary number n of such states are taken into account. See Fig. 4.2 for an overall picture for 4 stimuli, also with the episode states. Table 4.2 shows a conceptual matrix representation of the model depicted in conceptual graphical representation in Figs. 4.1 and 4.2. In Table 4.3 the connection weights are described.

The numerical representation of the model by a set of local dynamic properties involving differential equations is presented below and is summarized in Box 4.1. During processing, each state has a strength represented by a real number between 0 and 1. Parameter η is a speed factor, indicating the speed by which an activation level is updated upon received input from other states. Below, the dynamics are described in more detail subsequently for each state by a dynamic (temporally) Local Property (LP) specifying how the activation value for this state is updated (after a time step of Δt) based on the activation values of the states connected to it (the incoming arrows in Figs. 4.1 and 4.2).

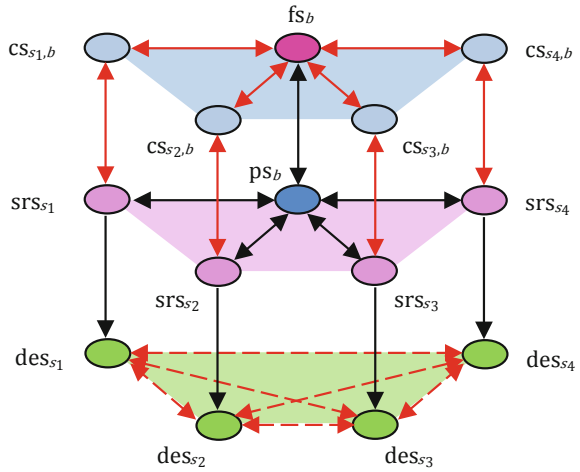
The numerical representation of the temporal-causal network model is discussed in the form of local dynamic properties LP1, ... for each of the states. In these dynamic local properties logistic sum combination functions are used. In the

Table 4.1 Overview of the states used

State	Explanation
ps_b	Preparation state for bodily response b
fs_b	Feeling state for b
srs_{s_k}	Sensory representation state for s_k
$cs_{s_k,b}$	Control state for regulation of sensory representation of s_k and feeling b
des_{s_k}	Dream episode state for s_k
mt_{s_k}	Memory trigger for s_k

Table 4.2 Conceptual matrix representation for the temporal-causal network model depicted in Figs. 4.1 and 4.2

To From	mt_{s_k}	srs_{s_k}	ps_b	$cs_{s_k,b}$	des_{s_k}	fs_b
mt_{s_k}		$\omega_{memory_triggering_s_k}$				
srs_{s_k}			$\omega_{responding_b}$	$\omega_{monitoring_s_k}$	$\omega_{manifesting_s_k}$	
ps_b		$\omega_{visualising_s_k}$				$\omega_{feeling_b}$
$cs_{s_k,b}$		$\omega_{suppressing_representation}$				$\omega_{suppressing_feeling}$
des_{s_k}					$\omega_{mutual_suppressing}$	
fs_b			$\omega_{amplifying_b}$	$\omega_{monitoring_b}$		
η_Y	—	$\eta_{srs_{s_k}}$	η_{ps_b}	$\eta_{cs_{s_k,b}}$	$\eta_{des_{s_k}}$	η_{fs_b}
$c_Y(\dots)$	—	$c_{srs_{s_k}}(\dots)$	$c_{ps_b}(\dots)$	$c_{cs_{s_k,b}}(\dots)$	$c_{des_{s_k}}(\dots)$	$c_{fs_b}(\dots)$

Fig. 4.2 Conceptual representation of the overall temporal-causal network model with four episode states and one feeling state ($m = 1$, $n = 4$)**Table 4.3** Overview of connections and weight names

From states	To state	Weight names	Connection types
$srs_{s_1}, \dots, srs_{s_n}$	ps_b	$\omega_{1,1}, \dots, \omega_{1,n}$	Responding
fs_b		ω_2	Amplifying
ps_b	fs_b	ω_3	Feeling
$cs_{s_1,b}, \dots, cs_{s_n,b}$	srs_{s_k}	$\omega_{4,1}, \dots, \omega_{4,n}$	Suppressing feeling
ps_b		$\omega_{5,k}$	Visualising
$cs_{s_k,b}$		$\omega_{6,k}$	Suppressing srs
mt_{s_k}		$\omega_{0,k}$	Memory triggering
srs_{s_k}	$cs_{s_k,b}$	$\omega_{7,k}$	Monitoring srs
fs_b		$\omega_{8,k}$	Monitoring feeling
srs_{s_k}	des_{s_k}	$\omega_{9,k}$	Episode manifesting
$des_{s_1}, \dots, des_{s_n}$		$\omega_{10,1,k}, \dots, \omega_{10,n,k}$	Mutual suppressing
$cs_{s_k,b}$		$\omega_{11,k}$	Traumatic suppressing

simulation experiments for the combination function in LP1–LP4 the following advanced logistic sum function is used:

$$\mathbf{alogistic}_{\sigma,\tau}(V_1, \dots, V_k) = \left(\frac{1}{1 + e^{-\sigma(V_1 + \dots + V_k - \tau)}} - \frac{1}{1 + e^{\sigma\tau}} \right) (1 + e^{-\sigma\tau})$$

Here σ is a steepness and τ a threshold parameter. In LP5 the simple logistic sum function is used:

$$\mathbf{slogistic}_{\sigma,\tau}(V_1, \dots, V_k) = \frac{1}{1 + e^{-\sigma(V_1 + \dots + V_k - \tau)}}$$

The first property LP1 describes how preparation for response b is affected by the sensory representations of stimuli s_k (triggering the response), and by the feeling state for b :

LP1 Preparation state for response b

$$\begin{aligned} \mathbf{dps}_b(t)/\mathbf{dt} &= \eta [\mathbf{alogistic}(\omega_{1,1} \mathbf{srs}_{s_1}(t), \dots, \omega_{1,n} \mathbf{srs}_{s_n}(t), \omega_2 \mathbf{fs}_b(t)) - \mathbf{ps}_b(t)] \\ \mathbf{ps}_b(t + \Delta t) &= \mathbf{ps}_b(t) + \eta [\mathbf{alogistic}(\omega_{1,1} \mathbf{srs}_{s_1}(t), \dots, \omega_{1,n} \mathbf{srs}_{s_n}(t), \omega_2 \mathbf{fs}_b(t)) - \mathbf{ps}_b(t)] \Delta t \end{aligned}$$

The feeling state for b is not only affected by a corresponding preparation state for b , but also by the (inhibiting) control states for s_k and b . This is expressed in dynamic property LP2. Note that for this suppressing effect the connection weight $\omega_{4,k}$ from the control state for s_k and b to feeling state for b is taken negative, for example $\omega_{4,k} = -1$.

LP2 Feeling state for b

$$\begin{aligned} \mathbf{d fs}_b(t)/\mathbf{dt} &= \eta [\mathbf{alogistic}(\omega_3 \mathbf{ps}_b(t), \omega_{4,1} \mathbf{cs}_{s_1, b}(t), \dots, \omega_{4,n} \mathbf{cs}_{s_n, b}(t)) - \mathbf{fs}_b(t)] \\ \mathbf{fs}_b(t + \Delta t) &= \mathbf{fs}_b(t) + \eta [\mathbf{alogistic}(\omega_3 \mathbf{ps}_b(t), \omega_{4,1} \mathbf{cs}_{s_1, b}(t), \dots, \omega_{4,n} \mathbf{cs}_{s_n, b}(t)) - \mathbf{fs}_b(t)] \Delta t \end{aligned}$$

The sensory representation state for s_k is triggered by memory state $\mathbf{mt}_{s_k}$ and further affected by the preparation state for b , and by the suppressing control state for s_k and b . For this suppressing effect the connection weight $\omega_{6,k}$ from the control state for s_k and b is taken negative. This is expressed in dynamic property LP3.

LP3 Sensory representation state for s_k

$$\begin{aligned} \mathbf{d srs}_{s_k}(t)/\mathbf{dt} &= \eta [\mathbf{alogistic}(\omega_{5k} \mathbf{ps}_b(t), \omega_{6,k} \mathbf{cs}_{s_k, b}(t), \omega_{6,k} \mathbf{mt}_{s_k}(t)) - \mathbf{srs}_{s_k}(t)] \\ \mathbf{srs}_{s_k}(t + \Delta t) &= \mathbf{srs}_{s_k}(t) + \eta [\mathbf{alogistic}(\omega_{5k} \mathbf{ps}_b(t), \omega_{6,k} \mathbf{cs}_{s_k, b}(t), \omega_{6,k} \mathbf{mt}_{s_k}(t)) - \mathbf{srs}_{s_k}(t)] \Delta t \end{aligned}$$

Note that property LP3 can be used to describe how the sensory representation of any traumatic s_k is triggered from memory, as a starting point for a dream: in a scenario the memory trigger values are taken 1. For non-traumatic s_k such triggering does not take place: the values are set to 0.

Activation of a control state for a specific sensory representation for s_k and b is based on the level of feeling b and the level of the sensory representation of s_k :

LP4 Control state for s_k and b

$$\begin{aligned} \mathbf{d cs}_{s_k, b}(t)/\mathbf{dt} &= \eta [\mathbf{alogistic}(\omega_{7,k} \mathbf{srs}_{s_k}(t), \omega_{8,k} \mathbf{fs}_b(t)) - \mathbf{cs}_{s_k, b}(t)] \\ \mathbf{cs}_{s_k, b}(t + \Delta t) &= \mathbf{cs}_{s_k, b}(t) + \eta [\mathbf{alogistic}(\omega_{7,k} \mathbf{srs}_{s_k}(t), \omega_{8,k} \mathbf{fs}_b(t)) - \mathbf{cs}_{s_k, b}(t)] \Delta t \end{aligned}$$

Due to the inherent parallelism in neural processes, at each point in time multiple sensory representation states can be active simultaneously. For cases of awake functioning the *Global Workspace Theory* (Baars 1997, 2002) was developed to describe how a single flow of conscious experience can come out of such a large multiplicity of (unconscious) processes; see also (Shanahan 2006) for an approach combining internal simulation and Global Workspace Theory, and (Dennett 1991, 2005) for a comparable perspective; see also (Windt and Noreika 2011).

The basic idea is that based on the various unconscious processes a *winner-takes-it-all competition* takes place to determine which one will get dominance and be included in the single flow of consciousness (after which it is accessible to all processes). This idea was applied in the dreaming context to determine which sensory representation elements will be included as an episode state des_{s_k} in a dream episode during the flow of dreams. This competition process is described in LP5, using inhibiting connections from the episode states des_{s_i} with $i \neq k$ to des_{s_k} . For the suppressing effects the connection weights from the des_{s_i} with $i \neq k$ to des_{s_k} are taken negative. Note that for the sake of notational simplicity $\omega_{10,k,k} = 0$ is used. For traumatic stimuli s_k an additional and strong way of inhibition of the corresponding episode state takes place, blocking the generation of an episode state for this stimulus. It is based on the control state for s_k and b and is assumed to have a strong negative connection strength $\omega_{11,k}$. For non-traumatic stimuli this connection is given strength 0; note that for the sake of simplicity this connection was not depicted in Fig. 4.2.

LP5 Episode state for s_k

$$\begin{aligned} \mathbf{d} \text{des}_{s_k}(t)/dt &= \eta [\text{slogistic}(\omega_{9k} \text{srs}_{s_k}(t), \omega_{11,k} \text{cs}_{s_k,b}(t), \omega_{10,1,k} \text{des}_{s_1}(t), \dots, \omega_{10,n,k} \\ \text{des}_{s_n}(t)) - \text{des}_{s_k}(t)] \\ \text{des}_{s_k}(t + \Delta t) &= \text{des}_{s_k}(t) + \eta [\text{slogistic}(\omega_{9k} \text{srs}_{s_k}(t), \omega_{11,k} \text{cs}_{s_k,b}(t), \omega_{10,1,k} \text{des}_{s_1}(t), \dots, \\ \omega_{10,n,k} \text{des}_{s_n}(t)) - \text{des}_{s_k}(t)] \Delta t \end{aligned}$$

LP1 Generating and amplifying a preparation state for response b

$$\begin{aligned} \mathbf{d} \text{ps}_b(t)/dt &= \eta [\text{alogistic}(\omega_{11} \text{srs}_{s_1}(t), \dots, \omega_{1n} \text{srs}_{s_n}(t), \omega_2 \text{fs}_b(t)) - \text{ps}_b(t)] \\ \text{ps}_b(t + \Delta t) &= \text{ps}_b(t) + \eta [\text{alogistic}(\omega_{11} \text{srs}_{s_1}(t), \dots, \omega_{1n} \text{srs}_{s_n}(t), \omega_2 \text{fs}_b(t)) - \text{ps}_b(t)] \Delta t \end{aligned}$$

LP2 Generating and regulating a feeling state for b

$$\begin{aligned} \mathbf{d} \text{fs}_b(t)/dt &= \eta [\text{alogistic}(\omega_3 \text{ps}_b(t), \omega_{41} \text{cs}_{s_1,b}(t), \dots, \omega_{4n} \text{cs}_{s_n,b}(t)) - \text{fs}_b(t)] \\ \text{fs}_b(t + \Delta t) &= \text{fs}_b(t) + \eta [\text{alogistic}(\omega_3 \text{ps}_b(t), \omega_{41} \text{cs}_{s_1,b}(t), \dots, \omega_{4n} \text{cs}_{s_n,b}(t)) - \text{fs}_b(t)] \Delta t \end{aligned}$$

LP3 Generating and regulating a sensory representation state for s_k

$$\begin{aligned} \mathbf{d} \text{srs}_{s_k}(t)/dt &= \eta [\text{alogistic}(\omega_{5k} \text{ps}_b(t), \omega_{6k} \text{cs}_{s_k,b}(t), \omega_{6k} \text{mt}_{s_k}(t)) - \text{srs}_{s_k}(t)] \\ \text{srs}_{s_k}(t + \Delta t) &= \text{srs}_{s_k}(t) + \eta [\text{alogistic}(\omega_{5k} \text{ps}_b(t), \omega_{6k} \text{cs}_{s_k,b}(t), \omega_{6k} \text{mt}_{s_k}(t)) - \text{srs}_{s_k}(t)] \Delta t \end{aligned}$$

LP4 Generating a control state for s_k and b

$$\begin{aligned} \mathbf{d} \text{cs}_{s_k,b}(t)/dt &= \eta [\text{alogistic}(\omega_{7k} \text{srs}_{s_k}(t), \omega_{8k} \text{fs}_b(t)) - \text{cs}_{s_k,b}(t)] \\ \text{cs}_{s_k,b}(t + \Delta t) &= \text{cs}_{s_k,b}(t) + \eta [\text{alogistic}(\omega_{7k} \text{srs}_{s_k}(t), \omega_{8k} \text{fs}_b(t)) - \text{cs}_{s_k,b}(t)] \Delta t \end{aligned}$$

LP5 Generating and suppressing an episode state for s_k

$$\begin{aligned} \mathbf{d} \text{des}_{s_k}(t)/dt &= \eta [\text{slogistic}(\omega_{9k} \text{srs}_{s_k}(t), \omega_{11,k} \text{cs}_{s_k,b}(t), \omega_{10,1,k} \text{des}_{s_1}(t), \dots, \omega_{10,n,k} \text{des}_{s_n}(t)) - \text{des}_{s_k}(t)] \\ \text{des}_{s_k}(t + \Delta t) &= \text{des}_{s_k}(t) + \eta [\text{slogistic}(\omega_{9k} \text{srs}_{s_k}(t), \omega_{11,k} \text{cs}_{s_k,b}(t), \omega_{10,1,k} \text{des}_{s_1}(t), \dots, \omega_{10,n,k} \text{des}_{s_n}(t)) - \text{des}_{s_k}(t)] \Delta t \end{aligned}$$

Box 4.1 The temporal-causal network model in numerical representation

4.4 Simulations of Example Dream Scenarios

A variety of simulation experiments have been performed, using numerical software. In the simulation experiments discussed below the settings were as shown in Table 4.3 (set by hand). As shown in the left hand side of the table, all non-inhibiting connections to preparation, feeling and control states have strength 1, and all inhibiting connections to feeling and sensory representation states have strengths -0.2 , resp. -0.5 , with an exception for the sensory representation state for s_1 , which is inhibited by strength -1 (due to a previous traumatic event involving s_1). Small differences in emotional association between the different s_k are expressed by different strengths from preparation of emotional response to sensory representation states, varying from 0.5 to 0.45. The sensory representation states are connected to the corresponding episode states with strength 1.2 and the latter states mutually inhibit each other by strength -0.6 . The threshold and steepness values used are shown in the right hand side of Table 4.4. Relatively low steepness values were used, except for the episode states. The threshold values for preparation and feeling states were taken 0.5; in order to model differences in emotional associations between the s_k , different threshold values were taken for their sensory representation and control states. The initial values of all states were set to 0, except for the initial value of srs_{s_1} which was set to 1 (a memory activation for a traumatic event). The speed factor η was 0.5, and the step size Δt was 0.1.

It may be convenient to read the scenario with a certain interpretation in mind. For example, s_1 may refer to a traumatic experience of seeing somebody who was dying (without having possibilities to save the person). Moreover, s_2 may refer to a situation where a presentation is due in a few minutes time, and no laptop nor slides

Table 4.4 Settings used for connection strength, threshold and steepness parameters: scenarios 1 and 2

From state	Connection		To state	Threshold	Steepness
srs _{s_k}	ω _{1,k}	1	ps _b	0.5	4
fs _b	ω ₂	1			
ps _b	ω ₃	1	fs _b	0.5	4
cs _{s_k,b}	ω _{4,k}	−0.2			
ps _b	ω _{5,1}	0.5	srs _{s₁}	0.5	4
cs _{s₁,b}	ω _{6,1}	−1			
ps _b	ω _{5,2}	0.5	srs _{s₂}	0.2	4
cs _{s₂,b}	ω _{6,2}	−0.5			
ps _b	ω _{5,3}	0.45	srs _{s₃}	0.22	4
cs _{s₃,b}	ω _{6,3}	−0.5			
srs _{s_k}	ω _{7,k}	1	cs _{s₁,b} cs _{s₂,b} cs _{s₃,b}	0.8	8
fs _b	ω _{8,k}	1		1.1	8
				1.4	8
srs _{s_k}	ω _{9,k}	1.2	es _{s_k}	0	200
es _{s_j}	ω _{10,j,k}	−0.6			

are available. Finally, s_3 may refer to a situation where an enormous traffic jam stands in the way of reaching an important meeting in time.

Scenario 1: no dream episode The first scenario discussed addresses the case where s_2 and s_3 do not play a role (by putting the connections from preparation to sensory representation $\omega_{5,k} = 0$ for $k \geq 2$). In Fig. 4.3 is shown that the control state for s_1 becomes active to reach level 0.2, and the activity of the sensory representation (indicated by rep) for s_1 drops to a low level. The preparation level stays below 0.5 and the feeling level stays below 0.4. Due to these modest values no dream episode state develops based on s_1 . The feeling level can be considered too low to seriously activate an internal simulation.

Scenario 2: two dream episodes In Fig. 4.4 a scenario is shown where the episode state des_{s_2} based on srs_{s_2} is succeeded (after time point 13) by an episode state des_{s_3} based on srs_{s_3} (see upper graph). Here the connection from preparation

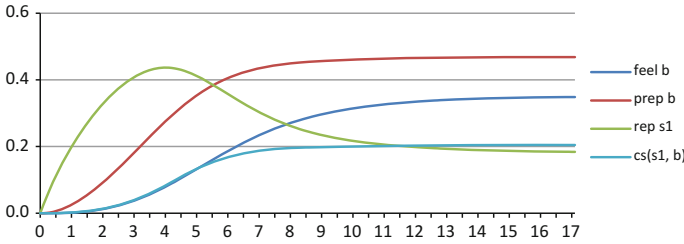


Fig. 4.3 Scenario 1: no strong feeling and no dream episode generated

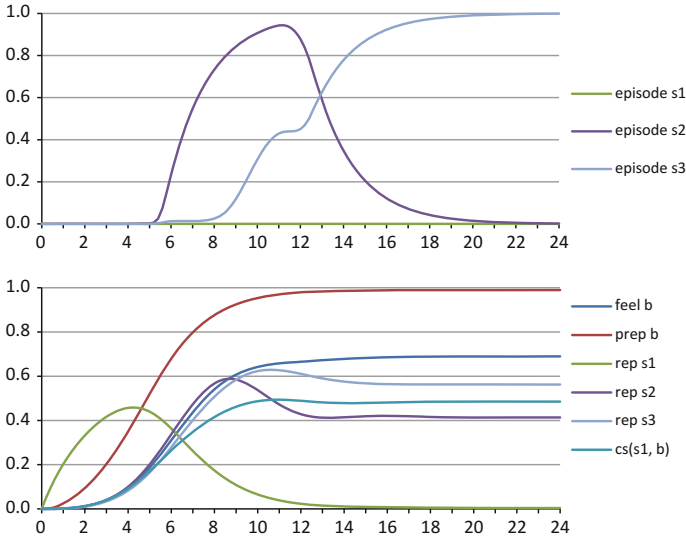


Fig. 4.4 Scenario 2: two subsequent dream episodes

for emotional response to sensory representation of s_3 has been given strength $\omega_{5,3} = 0.45$. As shown in the lower graph in Fig. 4.4, for this case the feeling level goes to 0.7, which is a situation in which regulation facilities become active. For example, due to this high feeling level the suppressing control state for s_1 becomes more active. In the lower graph of Fig. 4.4 the comparison between the sensory representations of s_2 and s_3 is shown; it is shown that first, up to time point 8, the sensory representation of s_2 dominates, reaching a level of around 0.6, which leads to a dream episode state des_{s_2} based on it, as shown in the upper graph in Fig. 4.4. But after time point 8 the sensory representation for s_2 is suppressed by the triggered regulation, and therefore beaten by the sensory representation for s_3 . As a consequence, after time point 13 the episode state for s_3 has won the competition, and provides the basis for a second dream episode. Note that the competition process took about 5 time units before the episode related to the sensory representation state that became the highest activated one at time 9 was able to beat the previous one.

Scenario 3: three dream episodes Similarly, scenarios for three or more dream episodes can be shown; see Fig. 4.5. Note that which episode states pop up depends on the association strengths to the emotional response. For example, if the emotional association strength $\omega_{5,3}$ for s_3 is made slightly lower, then the episode state for s_3 will never pop up due to the mutual inhibition. Moreover, the strength of the

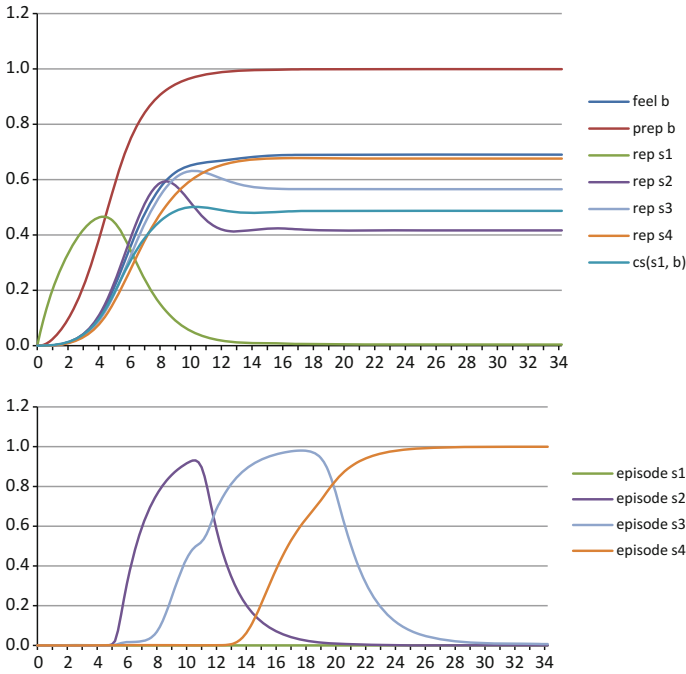


Fig. 4.5 Scenarion 3: Three subsequent dream episodes

Table 4.5 Settings used for connection strength, threshold and steepness parameters: scenario 3

from state	connection		to state	threshold	steepness
srs_{sk}	$\omega_{1,k}$	1	ps_b	0.5	4
fs_b	ω_2	1			
ps_b	ω_3	1	fs_b	0.5	4
$cs_{sk,b}$	$\omega_{4,k}$	-0.2			
ps_b	$\omega_{5,1}$	0.5	srs_{s1}	0.5	8
$cs_{s1,b}$	$\omega_{6,1}$	-2			
ps_b	$\omega_{5,2}$	0.5	srs_{s2}	0.25	8
$cs_{s2,b}$	$\omega_{6,2}$	-0.5			
ps_b	$\omega_{5,3}$	0.43	srs_{s3}	0.25	8
$cs_{s3,b}$	$\omega_{6,3}$	-0.5			
ps_b	$\omega_{5,4}$	0.38	srs_{s4}	0.25	8
$cs_{s4,b}$	$\omega_{6,4}$	-0.5			
srs_{sk} fs_b	$\omega_{7,k}$ $\omega_{8,k}$	1	$cs_{s1,b}$	0.7	8
			$cs_{s2,b}$	1.1	8
			$cs_{s3,b}$	1.4	8
			$cs_{s4,b}$	1.8	8
srs_{sk}	$\omega_{9,k}$	1	des_{sj}	0.25	60
des_{sj}	$\omega_{10,j,k}$	-0.4			

inhibition links affect whether or not two different episode states are considered compatible. If such mutual inhibition links have lower or zero strengths, then in one episode multiple (apparently compatible) episode states can co-occur.

In Table 4.5 the settings for this scenario are shown, where the cells with values that differ from the values in the previous scenarios are shaded in yellow.

4.5 Relations to Neurological Theories and Findings

In Levin and Nielsen (2007) dreaming is related to four main components Amygdala, Medial PreFrontal Cortex (MPFC), Hippocampus, Anterior Cingulate Cortex (ACC).

There is ample evidence of anatomical connections between the regions (..). Amygdala, in particular, is massively connected to the other regions in a reciprocal fashion (..). The four regions are also robustly connected to sensory, motor, and autonomic brain regions and thus are well suited to mediate higher cognitive functions, behaviors, and affective responses (Levin and Nielsen 2007, p. 505).

The biological counterparts of the preparation and sensory representation states in the model can be found in the sensory and (pre)motor cortices, indicated in Levin and Nielsen (2007) to be ‘robustly connected’ to the components as mentioned. The relations between sensory memory elements and their emotional associations are stored in the Hippocampus; in the model these relations are assumed to be fixed and modeled by the (bidirectional) connections between the sensory representations states s_{s_k} and preparation states ps_b of the emotional response b . The feeling state fs_b in the model can be related to the Amygdala, in combination with some limbic areas involved in maintaining ‘body maps’. As discussed in Sect. 4.2, the interaction between preparation state ps_b and feeling state fs_b is in line with the neurological theories of Damasio (1994, 1999, 2003, 2010). About the role of ACC empirical studies show evidence in different directions (e.g., Levin and Nielsen 2007, pp. 505–512); therefore it is not clear yet how it can be related to the model.

The interaction between MPFC and Amygdala has been extensively studied; e.g. Damasio (1994, 1999), Davidson (2002), Sotres-Bayon et al. (2004), Salzman and Fusi (2010), Levin and Nielsen (2007). In various empirical studies it has been found that lower activity of MPFC correlates to less controlled feeling levels, and, moreover, REM sleep is found to strengthen MPFC activation and reduce feeling levels. This regulating role of MPFC with respect to Amygdala activation makes these two neurological components suitable candidates for biological counterparts of the control state $cs_{s_k,b}$ and the feeling states fs_b in the temporal-causal network model presented in Sect. 4.3. As before, the connections between the two types of states may be related to the Hippocampus. Note that in the model the control states $cs_{s_k,b}$ also have a role in suppressing the activation of the corresponding sensory representation state s_{s_k} , which can be justified as being a form of emotion regulation by attentional deployment; e.g. Gross (1998, 2007); see also Sect. 4.2. The episode states des_{s_k} and their competition, as explained in Sect. 4.3 can be justified by referring to the Global Workspace Theory of consciousness by Baars (1997, 2002) and the perspective of Dennett (1991, 2005) who assume a similar competition mechanism.

4.6 Discussion

In this chapter based on a Network-Oriented Modeling approach, a temporal-causal network model was presented that models the generation of dream episodes from an internal simulation perspective, abstracting from a specific purpose. The contents of this chapter are based on (Treur 2011a).

The assumption that dreaming, especially when negative emotions are involved, can be considered as a purposeful form of internal simulation is widely supported; see, for example, for the purpose of improving coping skills to handle threatful situations (Revonsuo 2000; Valli et al. 2005; Valli and Revonsuo 2009), or for the purpose of strengthening fear emotion regulation capabilities (Levin and Nielsen

2007, 2009; Franzen et al. 2009; Gujar et al. 2011; Walker 2009; Walker and van der Helm 2009; Yoo et al. 2007; van der Helm et al. 2011; Rosales-Lagarde et al. 2012; Markarian et al. 2013; Mauss et al. 2013; Pace-Schott et al. 2015).

Internal simulation takes place when in response to sensory representation states activations associated preparation states for actions or bodily changes are activated, which in turn, by prediction links, activate other sensory representation states; see, for example, Hesslow (1994, 2002, 2012), Damasio (1994, 1999), Goldman (2006, Barsalou (2009), Marques and Holland (2009), Pezzulo et al. (2013). Building blocks to create such internal simulations are memory elements in the form of sensory representations and their associated emotions. The model exploits a mutual (winner-takes-it-all) competition process to determine sensory representation states that dominate in different dream episodes, comparable to one of the central ideas underlying Baars (1997, 2002)'s Global Workspace Theory of consciousness and Dennet (1991, 2005)'s perspective on consciousness. Emotion regulation mechanisms (Goldin et al. 2008; Gross 1998, 2007) were incorporated to regulate the activation levels of the feeling and the sensory representation states. The model was evaluated by a number of simulation experiments for scenarios with different numbers of dream episodes.

Note that the presented temporal-causal network model is meant as a plausible model of a human. Mechanisms identified in the neurological and cognitive literature were used in order to obtain a human-like temporal-causal network model, and to support its plausibility. Once such a human-like model is available, potential applications can be explored. A specific class of possible applications may concern virtual agents in the context of serious or non-serious gaming. In this context also some types of validation can be performed, for example, by evaluating how believable they are considered (also in dependence of parameter settings). Such applications and validations are a subject for future research.

A number of variations of the current model can be made. One variation is to take into account more than one emotion triggered by certain sensory representations. The model can easily be extended to cover this case. Another variation which is possible is to incorporate dependencies between sensory representations (e.g., resulting from sensory preconditioning; e.g. Brogden (1939), Hall (1996). As an extension of the current model, an adaptive temporal-causal network model has been developed for fear extinction learning during dreaming (Treur 2011b); this will be discussed in Chap. 5.

References

- B.J. Baars, *In the Theater of Consciousness: The Workspace of the Mind* (Oxford University Press, Oxford, 1997)
- B.J. Baars, The conscious access hypothesis: origins and recent evidence. *Trends Cogn. Sci.* **6**, 47–52 (2002)
- W. Barsalou, Simulation, situated conceptualization, and prediction. *Phil. Trans. R. Soc. B* **364**, 1281–1289 (2009)

- W. Becker, A.F. Fuchs, Prediction in the oculomotor system: smooth pursuit during transient disappearance of a visual target. *Exp. Brain Res.* **57**, 562–575 (1985)
- W.J. Brogden, Sensory pre-conditioning. *J. Exp. Psychol.* **25**, 323–332 (1939)
- A.R. Damasio, *Descartes' Error: Emotion, Reason and the Human Brain* (Papermac, London, 1994)
- A.R. Damasio, *The Feeling of What Happens. Body and Emotion in the Making of Consciousness* (Harcourt Brace, New York, 1999)
- A.R. Damasio, *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain* (Vintage books, London, 2003)
- A.R. Damasio, *Self Comes to Mind: Constructing the Conscious Brain* (Pantheon Books, NY, 2010)
- R.J. Davidson, Anxiety and affective style: role of prefrontal cortex and amygdala. *Biol. Psychiatry* **51**, 68–80 (2002)
- G. Deliens, M. Gilson, P. Peigneux, Sleep and the processing of emotions. *Exp. Brain. Res.* **232**, 1403–1414 (2014)
- D.C. Dennett, *Consciousness Explained* (Little Brown, 1991)
- D.C. Dennett, *Sweet Dreams: Philosophical Obstacles to a Science of Consciousness* (MIT press, Cambridge, MA, 2005)
- P.L. Franzen, D.J. Buysse, R.E. Dahl, W. Thompson, G.J. Siegle, Sleep deprivation alters pupillary reactivity to emotional stimuli in healthy young adults. *Biol. Psychol.* **80**, 300–305 (2009)
- P.R. Goldin, K. McRae, W. Ramel, J.J. Gross, The neural bases of emotion regulation: reappraisal and suppression of negative emotion. *Biol. Psychiatry* **63**, 577–586 (2008)
- A.I. Goldman, *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading* (Oxford University Press, New York, 2006)
- A.N. Goldstein, M.P. Walker, The role of sleep in emotional brain function. *Ann. Rev. Clin. Psychol.* **10**, 679–708 (2014)
- J.J. Gross, Antecedent- and response-focused emotion regulation: divergent consequences for experience, expression, and physiology. *J. Pers. Soc. Psychol.* **74**, 224–237 (1998)
- J.J. Gross, *Handbook of Emotion Regulation* (Guilford Press, New York, 2007)
- N. Gujar, S.A. McDonald, M. Nishida, M.P. Walker, A role for REM sleep in recalibrating the sensitivity of the human brain to specific emotions. *Cereb. Cortex* **21**, 115–123 (2011)
- G. Hall, Learning about associatively activated stimulus representations: implications for acquired equivalence and perceptual learning. *Anim. Learn. Behav.* **24**, 233–255 (1996)
- G. Hesslow, Will neuroscience explain consciousness? *J. Theoret. Biol.* **171**, 29–39 (1994)
- G. Hesslow, Conscious thought as simulation of behaviour and perception. *Trends Cogn. Sci.* **6**, 242–247 (2002)
- G. Hesslow, The current status of the simulation theory of cognition. *Brain Res.* **1428**, 71–79 (2012)
- J.A. Hobson, REM sleep and dreaming: towards a theory of protoconsciousness. *Nat. Rev. Neurosci.* **10**, 803–814 (2009)
- M. Kahn, G. Sheppes, A. Sadeh, Sleep and emotions: bidirectional links and underlying mechanisms. *Int. J. Psychophysiol.* **89**(2), 218–228 (2013)
- R. Levin, T.A. Nielsen, Disturbed dreaming, posttraumatic stress disorder, and affect distress: a review and neurocognitive model. *Psychol. Bull.* **133**, 482–528 (2007)
- R. Levin, T.A. Nielsen, Nightmares, bad dreams, and emotion dysregulation. A review and new neurocognitive model of dreaming. *Curr. Dir. Psychol. Sci.* **18**, 84–88 (2009)
- S.A. Markarian, S.M. Pickett, D.F. Deveson, B.B. Kanona, A model of BIS/BAS sensitivity, emotion regulation difficulties, and depression, anxiety, and stress symptoms in relation to sleep quality. *Psychiatry Res.* **210**(1), 281–286 (2013)
- H.G. Marques, O. Holland, *Architectures for functional imagination*. *Neurocomputing* **72**, 743–759 (2009)
- I.B. Mauss, A.S. Troy, M.K. LeBourgeois, Poorer sleep quality is associated with lower emotion-regulation ability in a laboratory paradigm. *Cogn. Emot.* **27**(3), 567–576 (2013)

- T.A. Nielsen, P. Stenstrom, What are the memory sources of dreaming? *Nature* **437**, 1286–1289 (2005)
- E.F. Pace-Schott, A. Germain, M.R. Milad, Effects of sleep on memory for conditioned fear and fear extinction. *Psychol. Bull.* **141**(4), 835–857 (2015)
- G. Pezzulo, M. Candidi, H. Dindo, L. Barca, Action simulation in the human brain: twelve questions. *New Ideas Psychol.* **31**, 270–290 (2013)
- A. Revonsuo, The reinterpretation of dreams: an evolutionary hypothesis of function of dreaming. *Behav. Brain Sci.* **23**, 877–901 (2000)
- A. Rosales-Lagarde, J.L. Armony, Y. del Río-Portilla, D. Trejo-Martínez, R. Conde, M. Corsi-Cabrera, Enhanced emotional reactivity after selective REM sleep deprivation in humans: an fMRI study. *Front. Behav. Neurosci.* **6**(25), 1–13 (2012)
- C.D. Salzman, S. Fusi, Emotion, cognition, and mental state representation in amygdala and prefrontal cortex. *Annu. Rev. Neurosci.* **33**, 173–202 (2010)
- M. Shanahan, A cognitive architecture that combines internal simulation with a global workspace. *Conscious. Cogn.* **15**, 433–449 (2006)
- F. Sotres-Bayon, D.E. Bush, J.E. LeDoux, Emotional perseveration: an update on prefrontal-amygdala interactions in fear extinction. *Learn. Mem.* **11**, 525–535 (2004)
- J. Treur, A computational agent model using internal simulation to generate emotional dream episodes, in *Proceedings of the Second International Conference on Biologically Inspired Cognitive Architectures, BICA'11*, ed. by A.V. Samsanovich, K.R. Johannsdottir (IOS Press, 2011a), pp. 389–399
- J. Treur, dreaming your fear away: a computational model for fear extinction learning during dreaming, in *Proceedings of the 18th International Conference on Neural Information Processing, ICONIP'11, Part III*. Lecture Notes in Artificial Intelligence, vol. 7064, ed. by B.-L. Lu, L. Zhang, J. Kwok (Springer, Berlin Heidelberg, 2011b), pp. 197–209
- K. Valli, A. Revonsuo, O. Palkas, K.H. Ismail, K.J. Ali, R.L. Punamaki, The threat simulation theory of the evolutionary function of dreaming: evidence from dreams of traumatized children. *Conscious. Cogn.* **14**, 188–218 (2005)
- K. Valli, A. Revonsuo, The threat simulation theory in light of recent empirical evidence: a review. *Am. J. Psychol.* **122**, 17–38 (2009)
- E. van der Helm, J. Yao, S. Dutt, V. Rao, J.M. Saletin, M.P. Walker, REM sleep depotentiates amygdala activity to previous emotional experiences. *Curr. Biol.* **21**(23), 1–4 (2011)
- M.P. Walker, The role of sleep in cognition and emotion. *Ann. N. Y. Acad. Sci.* **1156**, 168–197 (2009)
- M.P. Walker, E. van der Helm, Overnight therapy? The role of sleep in emotional brain processing. *Psychol. Bull.* **135**, 731–748 (2009)
- J.M. Windt, V. Noreika, How to integrate dreaming into a general theory of consciousness. *Conscious. Cogn.* **20**, 1091–1107 (2011). doi:[10.1016/j.concog.2010.09.010](https://doi.org/10.1016/j.concog.2010.09.010)
- S.S. Yoo, N. Gujar, P. Hu, F.A. Jolesz, M.P. Walker, The human emotional brain without sleep—a prefrontal amygdala disconnect. *Curr. Biol.* **17**, R877–R878 (2007)

Chapter 5

Dreaming Your Fear Away

Fear Extinction Learning During Dreaming

Abstract In this chapter, following a Network-Oriented Modeling approach, an adaptive temporal-causal network model is presented that models how dreaming is used to learn fear extinction. The network model addresses dreaming as internal simulation incorporating memory elements in the form of sensory representations and their associated fear. During dream episodes regulation of fear takes place, which is strengthened in an adaptive manner by Hebbian learning. The model was evaluated by a number of simulation experiments for different scenarios.

5.1 Introduction

In the recent cognitive and neurological literature the mechanisms and functions of dreaming have received much attention; e.g., (Hobson 2009; Levin and Nielsen 2007, 2009; Nielsen and Stenstrom 2005; Revonsuo 2000; Pace-Schott et al. 2015; van der Helm et al. 2011; Goldstein and Walker 2014; Deliens et al. 2014; Valli et al. 2005; Valli and Revonsuo 2009; Walker 2009; Walker and van der Helm 2009; Yoo et al. 2007).

In such literature, usually dreaming is considered a form of internal simulation of real-life-like processes serving as training in order to learn or adapt certain capabilities. Dreaming makes use of memory elements for sensory representations (mental images) and their associated emotions to generate ‘virtual simulations’; e.g., (Levin and Nielsen 2007, pp. 499–500). Taking into account fear emotions that often play an important role in dreams, strengthening of regulation of such emotions is considered an important purpose of dreaming; see, for example, (Levin and Nielsen 2007; Walker and van der Helm 2009; van der Helm et al. 2011; Deliens et al. 2014; Pace-Schott et al. 2015). To this end in dreams adequate exercising material is needed: sensory representations of emotion-loaded situations are activated, built on memory elements suitable for high levels of arousal:

They are recombined or remapped in order to introduce elements that are incompatible with existing fear memories, thus facilitating (among other functions) the acquisition or maintenance of extinction memories. The latter inhibit fear memories (..), and consequently alleviate affect load (Levin and Nielsen 2007, pp. 500–501).

A comparison can be made to a virtual reality form of exposure therapy (Levin and Nielsen 2007, pp. 500–501). Strong fear associations of the sensory memory elements used to make up a dream creates situations in which a person has to cope with high levels of fear.

Adopting basic elements from Chap. 4 the adaptive temporal-causal network model presented here generates the type of internal simulation that is assumed to take place in dreaming. For the different dream episodes, the internal simulation incorporates interrelated processes of activation of sensory representation states (from memory) providing mental images, and activation of associated feelings. Moreover, it incorporates emotion regulation to suppress the feeling levels and the sensory representation states. The regulation mechanism is adaptive by strengthening the relevant connections by Hebbian learning; e.g., (Bi and Poo 2001; Gerstner and Kistler 2002; Hebb 1949); this adaptive mechanism was described earlier in (Treur 2011b).

The structure of the chapter is as follows. In Sect. 5.2 the adaptive temporal-causal network model is described in more detail both in a conceptual and a numerical representation. Section 5.3 presents simulation results providing some dream scenarios. In Sect. 5.4 the relation of the model with neurological theories and findings is addressed. Finally, Sect. 5.5 is a discussion.

5.2 An Adaptive Temporal-Causal Network Model for Fear Extinction Learning

The adaptive temporal-causal network model presented here is based on mechanisms suggested in neurological literature. First a conceptual representation is discussed and next a numerical representation.

5.2.1 *Conceptual Representation of the Adaptive Network Model*

In Fig. 5.1 a conceptual graphical representation of the model is shown which gives an overview of the states and connections. Some of the (non-adaptive) basic elements were adopted from (Treur 2011a); see Chap. 4. In Fig. 5.1 the basic model for a given stimulus s_k with sensory representation state srs_{sk} and dream episode state des_{sk} is shown ($k = 1, \dots, n$). An explanation of the states used is shown in Table 5.1; an overview of the connections and their weights is shown in Table 5.2.

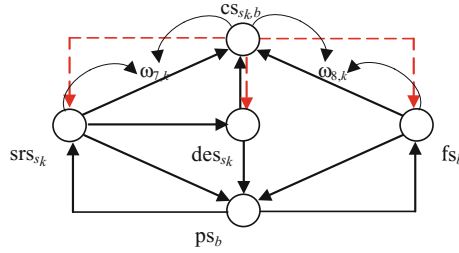


Fig. 5.1 Graphical conceptual representation of the adaptive temporal-causal network model

Table 5.1 Overview of the state variables used

State	Explanation
ps_b	Preparation state for bodily response b
fs_b	Feeling state for b
srs_{s_k}	Sensory representation state for stimulus s_k
$cs_{s_k,b}$	Control state for regulation of sensory representation of s_k and feeling b
des_{s_k}	Dream episode state for s_k
cs_{s_k}	Memory trigger for s_k

Table 5.2 Overview of connections and weights

From states	To state	Weights	LP
$srs_{s_1}, \dots, srs_{s_n}$ fs_b $des_{s_1}, \dots, des_{s_n}$	ps_b	$\omega_{1,1}, \dots, \omega_{1,n}$ ω_2 $\omega_{12,1}, \dots, \omega_{12,n}$	LP1
ps_b $cs_{s_1,b}, \dots, cs_{s_n,b}$	fs_b	ω_3 $\omega_{4,1}, \dots, \omega_{4,n}$	LP2
ps_b $cs_{s_k,b}$ mt_{s_k}	srs_{s_k}	$\omega_{5,k}$ $\omega_{6,k}$ $\omega_{0,k}$	LP3
srs_{s_k} fs_b des_{s_k}	$cs_{s_k,b}$	$\omega_{7,k}$ $\omega_{8,k}$ $\omega_{13,k}$	LP4
srs_{s_k} $des_{s_1}, \dots, des_{s_n}$ $cs_{s_k,b}$	des_{s_k}	$\omega_{9,k}$ $\omega_{10,1,k}, \dots, \omega_{10,n,k}$ $\omega_{11,k}$	LP5

Note that in Fig. 5.1 a sensory representation state and episode state for only one stimulus s_k is depicted. In the specification of the model below an arbitrary number n of such states are taken into account. In Sect. 5.3, a simulation scenario with four stimuli s_k is presented.

The inhibiting links for fear regulation are indicated by dotted arrows (in red) from $cs_{s_k,b}$ to srs_{s_k} , fs_b and des_{s_k} . The two links between srs_{s_k} and ps_b indicate the bidirectional association between stimulus representation srs_{s_k} and emotional

feeling fs_b . The links between ps_b and fs_b indicate a recursive as-if body loop (see below). The links from srs_{s_k} and $cs_{s_k,b}$ to $\omega_{7,k}$ and from fs_b and $cs_{s_k,b}$ to $\omega_{8,k}$ indicate the adaptive element in the model.

The adaptive network model incorporates four connected cycles (see Fig. 5.1):

- A positive preparation-feeling cycle $ps_b - fs_b$ (right lower part in Fig. 5.1)
- A positive preparation-sensory representation cycle $ps_b - srs_{s_k}$ (left lower part)
- A negative emotion regulation cycle $cs_{s_k,b} - fs_b, srs_{s_k}, des_{s_k}$ (upper part)
- A positive fear extinction learning cycle $cs_{s_k,b} - \omega_{7,k}, \omega_{8,k}$ (upper part)

Each of these cycles will be briefly discussed.

The preparation-feeling cycle $ps_b - fs_b$

As indicated in Sect. 5.1 above, dreams can be considered as flows of activated imaginations based on (re)combined sensory memory elements with emotional associations. Such flows can be related to the notion of internal simulation of body states put forward, among others, by (Damasio 1994, 1999). The idea of internal simulation of body states is that sensory representation states are activated (e.g., mental images), which in response trigger associated preparation states for actions or bodily changes, which, by prediction links, in turn activate sensory representation states for body states.

sensory representation states $\rightarrow$ preparation states
 $\rightarrow$ sensory representation states

Internal simulation has been used in general, for example, to describe prediction of effects of own actions (e.g., Becker and Fuchs 1985), processes in another person's mind (e.g., Goldman 2006) or processes in a person's own body (e.g., Damasio 1994). The idea of internal simulation is exploited in this cycle more in particular by applying it to bodily changes expressing emotions, using the notion of *as-if body loop* (e.g., Damasio 1994, pp. 155–158; Damasio 1999, pp. 79–80; Damasio 2010); see also Chap. 3:

sensory representation $\rightarrow$ preparation for bodily changes = emotional response $\rightarrow$
 emotion felt = based on sensory representation of (simulated) bodily changes

The preparation-sensory representation cycle $ps_b - srs_{s_k}$

Sensory representations as stored in memory usually have emotional responses associated to them. This means that as soon as a sensory representation is activated also its associated emotional response preparations are activated, and conversely, when an emotional response preparation is active, also the sensory representations associated to this type of response become active. This results in a cycle between sensory representations srs_{s_k} and emotional response preparations ps_b shown in the left lower part of Fig. 5.1. The link from ps_b to srs_{s_k} triggers a (predicted or expected) sensory representation upon activation of the (emotional) preparation

state ps_b , as a form of internal simulation. This can be related to, among others, (Hesslow 1994, 2002, 2012). Together with the preparation—feeling cycle discussed above, this provides a state of fear as a complex and cyclic activation state of fear response preparations, fear feelings and fearful sensory representations.

The emotion regulation cycle $cs_{s_k,b} - fs_b, srs_{s_k}, srs_{s_k}$

Fear extinction indicates the process of suppressing fear states. This can be considered a specific type of *emotion regulation* to control emotions that are felt as too strong; e.g., (Goldin et al. 2008; Gross 1998, 2007). Emotion regulation mechanisms cover *antecedent-focused regulation* (e.g., selection and modification of the situation, attentional deployment, and reappraisal) and *response-focused regulation* (suppression of a response). Regulation of high levels of fear can take place by antecedent-focused emotion regulation, for example, by *attentional deployment* in the form of focusing attention in such a way that situations or aspects of situations in which too strong fear-related stimuli occur are kept out of the attention focus, or by a form of *re-appraisal* decreasing the negative feeling level based on changing the cognitive interpretation of fear-related stimuli into a less negative one. In the upper part of Fig. 5.1 such an emotion regulation mechanism is depicted. The upward arrows to the control state $cs_{s_k,b}$ take care for monitoring the sensory representations srs_{s_k} , feeling state fs_b and dream episode state des_{s_k} for the fear state, and when the fear level is too high, this leads to activation of the relevant control states $cs_{s_k,b}$. These control states in turn lead to inhibition of the fear-related states (the downward, dotted arrows in the upper part of Fig. 5.1).

The fear extinction learning cycle $cs_{s_k,b} - \omega_{7,k}, \omega_{8,k}$

The basis of fear extinction learning is that the emotion regulation mechanisms discussed above are adaptive: they are strengthened over time when they are intensively used. Note that fear extinction learning is *not* a form of unlearning or extinction of acquired fear associations, but it is additional *learning of fear inhibition* in order to counterbalance the fear associations which themselves remain intact (e.g., Levin and Nielsen 2007, p. 507). This learning process is modeled by applying a Hebbian learning principle (e.g., Bi and Poo 2001; Gerstner and Kistler 2002; Hebb 1949) to the upward connections $\omega_{7,k}$ and $\omega_{8,k}$ from sensory representation state srs_{s_k} and feeling state fs_b to the control state $cs_{s_k,b}$ in the upper part of Fig. 5.1. Note that the dream episode state and its upward link to the control state serve as an amplifier in this Hebbian learning process. The positive cyclic character of this learning process is as follows: the stronger the upward connections become, the higher the activation level of the control state, and this again strengthens the learning process for the connections.

The adaptive temporal-causal network model has been designed according to the format discussed in Chap. 2. Parameter η is a speed factor, indicating the speed by which an activation level is updated upon received input from other states. During processing, each state has an activation level represented by a real number between 0 and 1.

5.2.2 Numerical Representation of the Adaptive Network Model

Below, the numerical representation of the model is discussed in the form of (temporally). Local Properties (LP) for the dynamics of the states based on the connections between the states; these local properties are described by differential equations.

In these specifications the advanced logistic sum combination function **alogistic** _{σ, τ} (...) is used with σ a steepness and τ a threshold parameter (see also Chap. 2):

$$\mathbf{alogistic}_{\sigma, \tau}(V_1, \dots, V_k) = \left(\frac{1}{1 + e^{-\sigma(V_1 + \dots + V_k - \tau)}} - \frac{1}{1 + e^{\sigma\tau}} \right) (1 + e^{-\sigma\tau})$$

Note that for higher values of $\sigma\tau$ (e.g., σ higher than $20/\tau$) this threshold function can be approximated by the simple logistic sum combination function **slogistic** _{σ, τ} ($V_1, \dots, V_k$); this has been used in LP5:

$$\mathbf{slogistic}_{\sigma, \tau}(V_1, \dots, V_k) = \frac{1}{1 + e^{-\sigma(V_1 + \dots + V_k - \tau)}}$$

The first property LP1 describes how preparation for response b is affected by the sensory representation and episode states of stimuli s_k (triggering the response), and by the feeling state for b :

LP1 Preparation state for response b

$$\frac{dps_b(t)}{dt} = \eta [\mathbf{alogistic}_{\sigma, \tau}(\omega_{1,1} srs_{s_1}(t), \dots, \omega_{1,k} srs_{s_k}(t), \omega_2 fs_b(t), \omega_{12,1} es_{s_1}(t), \dots, \omega_{12,k} es_{s_k}(t)) - ps_b(t)]$$

The feeling state for b is not only affected by a corresponding preparation state for b , but also by the inhibiting control states for s_k and b . This is expressed in dynamic property LP2. Note that for this suppressing effect the connection weight $\omega_{4,k}$ from the control state for s_k and b to feeling state for b is taken negative, for example $\omega_{4,k} = -1$.

LP2 Feeling state for b

$$\frac{dfs_b(t)}{dt} = \eta [\mathbf{alogistic}_{\sigma, \tau}(\omega_3 ps_b(t), \omega_{4,1} cs_{s_1, b}(t), \dots, \omega_{4,k} cs_{s_k, b}(t)) - fs_b(t)]$$

The sensory representation state for s_k is affected by the preparation state for b (fear association) and by the suppressing control state for s_k and b . For this suppressing effect the connection weight $\omega_{6,k}$ from the control state for s_k and b is taken negative. This is expressed in dynamic property LP3. Moreover, property LP3 is used to describe how the sensory representation of any traumatic s_k is triggered from memory, as a starting point for a dream: in a scenario the memory trigger values are set 1. For non-traumatic s_k such triggering does not take place: the values are set 0.

LP3 Sensory representation state for s_k

$$dsrs_{s_k}(t)/dt = \eta [\text{alogistic}_{\sigma,\tau}(\omega_{5,k} ps_b(t), \omega_{6,k} cs_{s_k,b}(t), \omega_{0,k} mt_{s_k}(t)) - srs_{s_k}(t)]$$

Activation of a control state for a specific sensory representation for s_k and b is based on the level for feeling b and the activation level of the sensory representation and episode states of s_k :

LP4 Control state for s_k and b

$$dcs_{s_k,b}(t)/dt = \eta [\text{alogistic}_{\sigma,\tau}(\omega_{7,k} srs_{s_k}(t), \omega_{8,k} fs_b(t), \omega_{13,k} des_{s_k}(t)) - cs_{s_k,b}(t)]$$

Due to the inherent parallelism in neural processes, at each point in time multiple sensory representation states can be active simultaneously. For cases of awake functioning the *Global Workspace Theory* (e.g., Baars 1997) was developed to describe how a single flow of conscious experience can come out of such a large multiplicity of (unconscious) parallel processes. The basic idea is that based on the various unconscious processes a *winner-takes-it-all competition* takes place to determine which one will get dominance and be included in the single flow of consciousness. This idea was applied here in the dreaming context to determine which sensory representation element will be included as an episode state des_{s_k} in a dream. This competition process is described in LP5, using mutual inhibiting connections from episode states des_{s_i} with $i \neq k$ to des_{s_k} . For the suppressing effects the connection weights from the des_{s_i} with $i \neq k$ to des_{s_k} are taken negative, for example $\omega_{10,i,k} = -0.6$ for $i \neq k$. Note that for the sake of notational simplicity $\omega_{10,k,k} = 0$ is taken. For traumatic stimuli s_k an additional and strong way of inhibition of the corresponding episode state takes place, blocking the generation of an episode state for this stimulus. It is based on the control state for s_k and b and is assumed to have a strong negative connection strength $\omega_{11,k}$. For non-traumatic stimuli this connection is given strength 0.

LP5 Episode state for s_k

$$des_{s_k}(t)/dt = \eta [\text{slogistic}_{\sigma,\tau}(\omega_{9,k} srs_{s_k}(t), \omega_{11,k} cs_{s_k,b}(t), \omega_{10,1,k} des_{s_1}(t), \dots, \omega_{10,i,k} des_{s_i}(t)) - des_{s_k}(t)]$$

Hebbian learning to strengthen fear extinction

From a Hebbian perspective, strengthening of a connection over time may take place when both nodes are often active simultaneously ('neurons that fire together wire together'). The principle goes back to Hebb (1949), but has recently gained enhanced interest by more extensive empirical support (e.g., Bi and Poo 2001), and more advanced mathematical formulations (e.g., Gerstner and Kistler 2002); also see Chap. 2, Sect. 2.10. In the adaptive temporal-causal network model two upward connections that play a role in monitoring for the emotion regulation cycle are

adapted based on a Hebbian learning principle. More specifically, for such a connection from state X_i to state X_j its strength ω_{ij} is adapted using the following *Hebbian learning rule*, taking into account a maximal connection strength 1, a *learning rate* $\eta_{\omega_{ij}}$, and an *extinction rate* $\zeta_{\omega_{ij}}$, usually taken small (see also Chap. 2, Sect. 2.10):

$$\begin{aligned} d\omega_{ij}(t)/dt &= \eta_{\omega_{ij}} X_i(t) X_j(t) (1 - \omega_{ij}(t)) - \zeta_{\omega_{ij}} \omega_{ij}(t) \\ &= \eta_{\omega_{ij}} X_i(t) X_j(t) - (\eta_{\omega_{ij}} X_i(t) X_j(t) + \zeta_{\omega_{ij}}) \omega_{ij}(t) \end{aligned}$$

Here $X_i(t)$ and $X_j(t)$ are the activation levels of state X_i and X_j at time t and $\omega_{ij}(t)$ is the strength of the connection from state X_i to state X_j at time t . A similar Hebbian learning rule can be found in (Gerstner and Kistler 2002, p. 406). This learning principle has been applied (simultaneously) to the two upward connections from sensory representation and feeling states to the control state in Fig. 5.1, according to the following instantiations of the general learning rule above:

$$\begin{aligned} d\omega_{7,k}(t)/dt &= \eta_{\omega_{7,k}} \text{srs}_{s_k}(t) \text{cs}_{s_k,b}(t) (1 - \omega_{7,k}(t)) - \zeta_{\omega_{7,k}} \omega_{7,k}(t) \\ &= \eta_{\omega_{7,k}} \text{srs}_{s_k}(t) \text{cs}_{s_k,b}(t) - (\eta_{\omega_{7,k}} \text{srs}_{s_k}(t) \text{cs}_{s_k,b}(t) + \zeta_{\omega_{7,k}}) \omega_{7,k}(t) \\ d\omega_{8,k}(t)/dt &= \eta_{\omega_{8,k}} \text{fs}_b(t) \text{cs}_{s_k,b}(t) (1 - \omega_{8,k}(t)) - \zeta_{\omega_{8,k}} \omega_{8,k}(t) \\ &= \eta_{\omega_{8,k}} \text{fs}_b(t) \text{cs}_{s_k,b}(t) - (\eta_{\omega_{8,k}} \text{fs}_b(t) \text{cs}_{s_k,b}(t) + \zeta_{\omega_{8,k}}) \omega_{8,k}(t) \end{aligned}$$

In principle, the learning rate η and extinction rate ζ , can be taken differently for the different connections. In the example simulations discussed in Sect. 5.3 (shown in Fig. 5.2) the following values have been used: $\eta = 0.7$ for all $\omega_{7,k}$ and $\eta = 0.4$ for all $\omega_{8,k}$, and $\zeta = 0.001$ for all $\omega_{7,k}$ and $\omega_{8,k}$.

5.3 Simulations of Fear Extinction Learning in Dream Scenarios

In the dream scenarios that are described in this section the cycles as discussed play their roles as follows.

Triggering s_1

- A stimulus s_1 is given for which previously a high extent of fear has been experienced, and for which from time to time (in particular during sleep) a sensory representation state is triggered by memory (for the model this is considered an external trigger); note that such a memory trigger was not used for the other stimuli: their activation automatically happens due to the high fear levels induced by triggering s_1 , and maintained by the subsequent dream episodes.

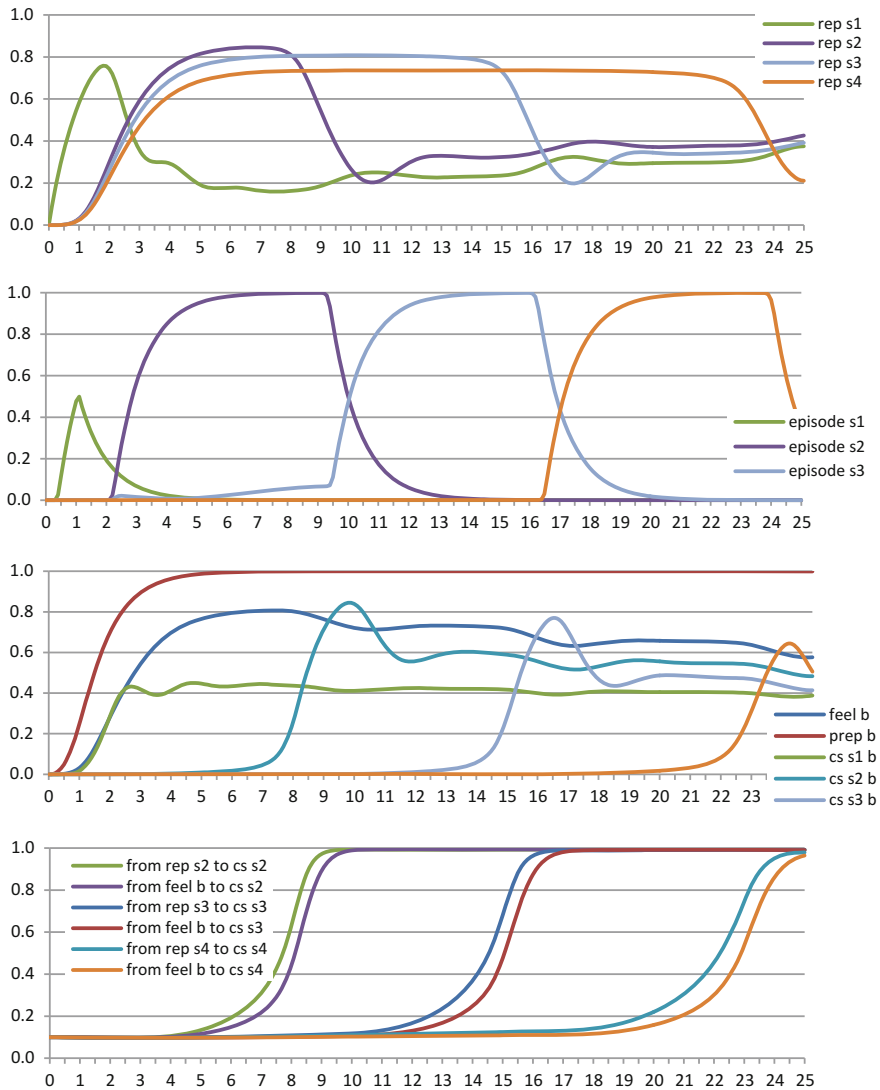


Fig. 5.2 Dream with three episodes showing extinction learning and reduction of feeling level

- The activation of the sensory representation of s_1 leads to activation of an enhanced preparation level for a bodily fear response b

The positive preparation-feeling cycle $ps_b - fs_b$

- By an as-if body loop an enhanced preparation level for b leads to an enhanced fear feeling level for b and vice versa

Blocking s_1

- By a strong form of emotion regulation in particular the sensory representation and episode state of s_1 are strongly suppressed: the activation level of the sensory representation of s_1 becomes low, and no dream episode state for s_1 occurs, as this is blocked

The positive preparation-sensory representation cycle $ps_b - srs_{s_k}$

- Other fear-associated stimuli s_k for $k \geq 2$ are available for which the person has less strong previous experiences; the sensory representation states for these s_k are activated by links from the high preparation state for b , depending on the strength of these links
- When the sensory representation state of a stimulus s_k is activated, this leads to an enhanced activation level of the preparation state for the emotional fear response

The positive preparation-feeling cycle $ps_b - fs_b$

- Due to the higher activation level of preparation for fear based on b , via the as-if body loop also the feeling level for b becomes higher: the person experiences more fear

Competition to achieve a dream episode des_{s_k}

- The active sensory representation for some s_k leads to a corresponding dream episode state, according to a competition process by mutual inhibition to get dominance in the episode

The negative emotion regulation cycle $cs_{s_k,b} - fs_b, srs_{s_k}, des_{s_k}$

- By the control states for emotion regulation for an active sensory representation for s_k both the fear feeling level and the sensory activation level of s_k are suppressed (attentional deployment)

The fear extinction learning cycle $cs_{s_k,b} - \omega_{7,k}, \omega_{8,k}$

- Due to nonzero activation levels of the control states and the fear feeling state for b , and the sensory representation and episode states for s_k Hebbian learning takes place strengthening the connections from feeling state and sensory representation to control state
- Increased connection strengths lead to higher activation levels for the control states

A variety of simulation experiments have been performed according to such scenarios, using numerical software. In the experiments discussed below (see Fig. 5.2) the settings were as shown in Table 5.3.

As shown in the left hand side of the table, all non-inhibiting connections to preparation, feeling, control, and episode states have strength 1, and the inhibiting connections from control states to feeling, sensory representation states and episode

Table 5.3 Settings used for connection strength, threshold and steepness parameters

From state	Connection		To state	Threshold	Steepness
srs_{s_k}	$\omega_{1,k}$	1	ps_b	0.5	4
fs_b	ω_2	1			
es_{s_k}	$\omega_{12,k}$	1			
ps_b	ω_3	1	fs_b	0.5	4
$cs_{s_k,b}$	$\omega_{4,k}$	-0.2			
ps_b	$\omega_{5,1}$	0.5	srs_{s_1}	0.25	8
$cs_{s_4,b}$	$\omega_{6,1}$	-2			
ps_b	$\omega_{5,2}$	0.5	srs_{s_2}	0.25	8
$cs_{s_2,b}$	$\omega_{6,2}$	-0.5			
ps_b	$\omega_{5,3}$	0.45	srs_{s_3}	0.25	8
$cs_{s_3,b}$	$\omega_{6,3}$	-0.5			
ps_b	$\omega_{5,4}$	0.4	srs_{s_4}	0.25	8
$cs_{s_4,b}$	$\omega_{6,4}$	-0.5			
srs_{s_1}	$\omega_{7,1}$	1	$cs_{s_k,b}$	1	8
fs_b	$\omega_{8,1}$	1			
es_{s_k}	$\omega_{13,k}$	0.3			
srs_{s_k}	$\omega_{9,k}$	1	es_{s_k}	0.25	60
$es_{s_j} (j \neq k)$	$\omega_{10,j,k}$	-0.6			
$es_{s_k,b}$ ($k \geq 2$)	$\omega_{11,k}$	-0.2			
$cs_{s_1,b}$	$\omega_{11,1}$	-20			

states, and mutually between episode states have strengths -0.2 , -0.5 , -0.2 , and -0.6 , respectively, with an exception for the sensory representation and episode states for s_1 , which are inhibited by strength -2 and -20 (they are blocked due to a previous traumatic event involving s_1). Small differences in emotional associations for the different s_k are expressed by different strengths from preparation of emotional response to sensory representation states, varying from 0.5 to 0.4. In the scenarios considered, the memory trigger for the sensory representation of s_1 has level 1 and connection strength 0.5. The threshold and steepness values used are shown in the right hand side of Table 5.3. Relatively low steepness values were used, except for the episode states. The threshold values for preparation and feeling states were taken 0.5; in order to model differences in emotional associations between the s_k , different threshold values were taken for their sensory representation and control states. The initial values of all states were taken 0, and for the adaptive connection strengths 0.1 initially (which also could be taken 0). The speed factor η was 1, and the step size Δt used was 0.1. For learning and extinction rates the following values have been used: $\eta_{\omega_{7,k}} = 0.7$ for all $\omega_{7,k}$ and $\eta_{\omega_{8,k}} = 0.4$ for all $\omega_{8,k}$, and $\zeta_{\omega_{7,k}} = \zeta_{\omega_{8,k}} = 0.001$ for all $\omega_{7,k}$ and $\omega_{8,k}$.

The example scenario discussed addresses a case where three dream episodes occur, related to the sensory representations of s_2 , s_3 , s_4 , subsequently. In Fig. 5.2 time is on the horizontal axis and the activation levels of the indicated states and connections are on the vertical axis. In the first graph it is shown that right from the start the sensory representation for s_1 becomes active (triggered from memory). Immediately the emotional response preparation for b starts to develop, and the related feeling, as shown in the third graph. Also in the third graph it is shown how as a result the control state for s_1 becomes active. Due to the strong suppression, no (full) dream episode develops for s_1 , as shown in the second graph. Due to the relatively high emotional response and feeling level, the sensory representations for s_2 , s_3 , s_4 become active, following that order and strength (first graph).

In a cyclic process, this further increases the emotional response preparation and feeling levels (third graph). As the sensory representation of s_2 is the strongest, it wins the competition for the dream episode from time point 3–9 (second graph).

Given this first episode and the high feeling and sensory representation levels, extinction learning takes place of the connections to the control state for s_2 (see fourth graph), reaching strengths one around 1 at time point 9, and hand in hand with this process the level of the control state for s_2 jumps up from time point 7 on (see third graph). As a result of this, control is exerted, suppressing after time point 9 the feeling level (third graph), the sensory representation of s_2 (first graph), and the related episode (second graph). As the feeling level was only partly reduced, and the sensory representation for s_2 does not compete anymore, from time point 11 on a second episode occurs, based on the sensory representation of s_3 (second graph). Again the whole adaptation process occurs, this time related to s_3 . From time point 16 on, this brings the feeling level more down (third graph), and suppresses the sensory representation of s_3 (first graph), and the related episode (second graph). After this, the whole process repeats itself for a third dream episode, based on the sensory representation of s_4 . This leads to another reduction of the feeling level around time 25. Overall, all connections for fear extinction in relation to the most strongly fear-related sensory representations have been learned and have values around 1, and the feeling level was reduced to below 0.6.

5.4 Relating the Adaptive Temporal-Causal Network Model to Neurological Theories

Recall from Chap. 4 that according to (Levin and Nielsen 2007) dreaming is related to four main brain components and their connections: Amygdala, Medial Prefrontal Cortex (MPFC), Hippocampus, Anterior Cingulate Cortex (ACC). Note that the biological counterparts of the preparation and sensory representation states in the model can be found in the sensory and (pre)motor cortices, indicated in (Levin and Nielsen 2007, p. 505) to be ‘robustly connected’ to the above mentioned components. One of the roles of the Hippocampus is to store and maintain the relations

between sensory memory elements and their emotional associations; in the model these connections are assumed to be fixed and modeled by the (bidirectional) connections between the sensory representations states srs_{s_k} and preparation states ps_b of the emotional response b . The feeling state fs_b in the model can be related to the Amygdala, possibly in combination with some limbic areas involved in maintaining ‘body maps’. As discussed in Sect. 5.2, the interaction between preparation state ps_b and feeling state fs_b is in line with the neurological theories of Damasio (1994, 1999, 2003, 2010). About the role of ACC empirical studies show evidence in different directions (e.g., Levin and Nielsen 2007, pp. 505–512); therefore it is not clear yet what exactly its function is in dreaming and how it can be related to the model presented in Sect. 5.2.

Especially the interaction between MPFC and Amygdala in fear extinction during dreaming has been extensively studied; (e.g. Davidson 2002; Levin and Nielsen 2007; Salzman and Fusi 2010; Sotres-Bayon et al. 2004). In various empirical studies it has been found that lower activity of MPFC correlates to less controlled feeling levels, and, moreover, REM sleep is found to strengthen MPFC activation and reduce feeling levels; see, for example, (Goldin et al. 2008; Gujar et al. 2011; Levin and Nielsen 2007; Walker and van der Helm 2009; Yoo et al. 2007). This regulating role of MPFC with respect to Amygdala activation makes MPFC a suitable candidate for biological counterpart of the control state $cs_{s_k,b}$ in the temporal-causal network model presented in Sect. 5.2. Moreover, the reported finding suggests that fear extinction learning affects activation of MPFC; this is in accordance with the modeling choice that the Hebbian learning was applied to the two upward connections from sensory representation and feeling states to the control state. As before, the connections between the two types of states may be related to the Hippocampus. Note that in the temporal-causal network model the control states $cs_{s_k,b}$ also have a role in suppressing the activation of the corresponding sensory representation state srs_{s_k} which can be justified as being a form of emotion regulation by attentional deployment (Gross 1998, 2007); see also Sect. 5.2. The episode states des_{s_k} and their competition can be justified by referring to the Global Workspace Theory of consciousness (e.g., Baars 1997), as explained in Sect. 5.2.

5.5 Discussion

In this chapter, following a Network-Oriented Modeling approach an adaptive temporal-causal network model was presented that models the generation of dream episodes from an internal simulation perspective, and uses these episodes for fear extinction learning. The contents of this chapter are based on (Treur 2011b).

The assumption that dreaming, especially when negative emotions are involved, can be considered as a purposeful form of internal simulation is widely supported, in particular, for the purpose of strengthening fear emotion regulation capabilities; e.g., (Levin and Nielsen 2007; Walker and van der Helm 2009; van der Helm et al.

2011; Deliens et al. 2014; Pace-Schott et al. 2015). Building blocks to create such internal simulations are memory elements in the form of sensory representations and their associated emotions. The model exploits a mutual (winner-takes-it-all) competition process to determine sensory representation states that dominate in different dream episodes, comparable to one of the central ideas underlying the Global Workspace Theory of consciousness (Baars 1997). Adaptive emotion regulation mechanisms (e.g., Bi and Poo 2001; Gerstner and Kistler 2002; Hebb 1949) were incorporated to regulate the activation levels of the feeling (by re-appraisal) and the sensory representation states (by attentional deployment). Adaptation in the model is based on Hebbian learning (see also Chap. 2, Sect. 2.10). The temporal-causal network model was evaluated by a number of simulation experiments for scenarios with different numbers of dream episodes.

References

- B.J. Baars, *In the Theater of Consciousness: The Workspace of the Mind* (Oxford University Press, Oxford, 1997)
- W. Becker, A.F. Fuchs, Prediction in the oculomotor system: smooth pursuit during transient disappearance of a visual target. *Exp. Brain Res.* **57**, 562–575 (1985)
- G. Bi, M. Poo, Synaptic modification by correlated activity: Hebb's postulate revisited. *Annu. Rev. Neurosci.* **24**, 139–166 (2001)
- A.R. Damasio, *Descartes' Error: Emotion, Reason and the Human Brain* (Papermac, London, 1994)
- A.R. Damasio, *The Feeling of What Happens. Body and Emotion in the Making of Consciousness* (Harcourt Brace, New York, 1999)
- A.R. Damasio, *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain* (Vintage books, London, 2003)
- A.R. Damasio, *Self Comes to Mind: Constructing the Conscious Brain* (Pantheon Books, NY, 2010)
- R.J. Davidson, Anxiety and affective style: role of prefrontal cortex and amygdala. *Biol. Psychiatry* **51**, 68–80 (2002)
- G. Deliens, M. Gilson, P. Peigneux, Sleep and the processing of emotions. *Exp. Brain Res.* **232**: 1403–1414 (2014)
- W. Gerstner, W.M. Kistler, Mathematical formulations of Hebbian learning. *Biol. Cybern.* **87**, 404–415 (2002)
- P.R. Goldin, K. McRae, W. Ramel, J.J. Gross, The neural bases of emotion regulation: reappraisal and suppression of negative emotion. *Biol. Psychiatry* **63**, 577–586 (2008)
- A.I. Goldman, *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading* (Oxford Univ. Press, New York, 2006)
- A.N. Goldstein, M.P. Walker, The role of sleep in emotional brain function. *Ann. Rev. Clin. Psychol.* **10**, 679–708 (2014)
- J.J. Gross, Antecedent- and response-focused emotion regulation: divergent consequences for experience, expression, and physiology. *J. Personal. Soc. Psych.* **74**, 224–237 (1998)
- J.J. Gross, *Handbook of Emotion Regulation* (Guilford Press, New York, 2007)
- N. Gujar, S.A. McDonald, M. Nishida, M.P. Walker, A role for REM sleep in recalibrating the sensitivity of the human brain to specific emotions. *Cerebral Cortex.* **21**, 115–123 (2011)
- D. Hebb, *The Organisation of Behavior* (Wiley, 1949)
- G. Hesslow, Will neuroscience explain consciousness? *J. Theoret. Biol.* **171**, 29–39 (1994)

- G. Hesslow, Conscious thought as simulation of behaviour and perception. *Trends Cogn. Sci.* **6**, 242–247 (2002)
- G. Hesslow, The current status of the simulation theory of cognition. *Brain Res.* **1428**, 71–79 (2012)
- J.A. Hobson, REM sleep and dreaming: towards a theory of protoconsciousness. *Nat. Rev. Neurosci.* **10**, 803–814 (2009)
- R. Levin, T.A. Nielsen, Disturbed dreaming, posttraumatic stress disorder, and affect distress: a review and neurocognitive model. *Psychol. Bull.* **133**, 482–528 (2007)
- R. Levin, T.A. Nielsen, Nightmares, bad dreams, and emotion dysregulation. A review and new neurocognitive model of dreaming. *Curr. Dir. Psychol. Sci.* **18**, 84–88 (2009)
- T.A. Nielsen, P. Stenstrom, What are the memory sources of dreaming? *Nature* **437**, 1286–1289 (2005)
- E.F. Pace-Schott, A. Germain, M.R. Milad, Effects of sleep on memory for conditioned fear and fear extinction. *Psychol. Bull.* **141**(4), 835–857 (2015)
- A. Revonsuo, The reinterpretation of dreams: an evolutionary hypothesis of function of dreaming. *Beh. Brain Sci.* **23**, 877–901 (2000)
- C.D. Salzman, S. Fusi, Emotion, cognition, and mental state representation in amygdala and prefrontal cortex. *Annu. Rev. Neurosci.* **33**, 173–202 (2010)
- F. Sotres-Bayon, D.E. Bush, J.E. LeDoux, Emotional perseveration: an update on prefrontal-amygdala interactions in fear extinction. *Learn. Mem.* **11**, 525–535 (2004)
- J. Treur, A computational agent model using internal simulation to generate emotional dream episodes, in *Proceedings of the Second Intern. Conf. on Biologically Inspired Cognitive Architectures*, ed. by A.V. Samsonovich et al. (BICA'11. IOS Press, in press, 2011a)
- J. Treur, Dreaming your fear away: a computational model for fear extinction learning during dreaming, in *Proceedings of the 18th International Conference on Neural Information Processing, ICONIP'11, Part III*, vol. 7064, ed. by B. -L. Lu, L. Zhang, J. Kwok. *Lecture Notes in Artificial Intelligence* (Springer, Berlin, Heidelberg, 2011b), pp. 197–209
- K. Valli, A. Revonsuo, The threat simulation theory in light of recent empirical evidence: a review. *Am. J. Psychol.* **122**, 17–38 (2009)
- K. Valli, A. Revonsuo, O. Palkas, K.H. Ismail, K.J. Ali, R.L. Punamaki, The threat simulation theory of the evolutionary function of dreaming: evidence from dreams of traumatized children. *Conscious Cogn.* **14**, 188–218 (2005)
- E. van der Helm, J. Yao, S. Dutt, V. Rao, J.M. Saletin, M.P. Walker, REM sleep depotentiates amygdala activity to previous emotional experiences. *Curr. Biol.* **21**(23), 1–4 (2011)
- M.P. Walker, The role of sleep in cognition and emotion. *Ann. NY Acad. Sci.* **1156**, 168–197 (2009)
- M.P. Walker, E. van der Helm, Overnight therapy? The role of sleep in emotional brain processing. *Psychol. Bull.* **135**, 731–748 (2009)
- S.S. Yoo, N. Gujar, P. Hu, F.A. Jolesz, M.P. Walker, The human emotional brain without sleep—a prefrontal amygdala disconnect. *Curr. Biol.* **17**, R877–R878 (2007)

Chapter 6

Emotions as a Vehicle for Rationality in Decision Making

Experiencing Emotions for Decisions Based on Experience

Abstract In this chapter, following a Network-Oriented Modeling approach, an adaptive temporal-causal network model is presented for decision making using valuing of action options based on predictive loops through feeling states. Hebbian learning is used for different types of connections in the adaptive network model. To assess the extent of rationality, a measure is introduced reflecting the environment's behaviour. Simulation results and the extents of rationality of the different models over time are presented and analysed.

6.1 Introduction

Decision making has often been considered as a rational process in which emotions can only play a disturbing role. In more recent times this has been questioned. For example, Loewenstein and Lerner (2003, p. 619) point at the positive functions served by emotions (see also Chap. 3, Sect. 3.5, and the quote in Chap. 1, Sect. 1.2). In particular, in decision making it may be questioned whether you can make an adequate decision without feeling good about it. Decisions with bad feelings associated to them may lack robustness. Many occasions may occur over time that trigger a temptation to change it into a decision with a better associated feeling. Decision making usually considers a number of options for an action to be chosen. Such a choice is often based on some form of valuing the options. In this valuing process emotions come in: some of the options relate to a more positive feeling than the other options. It has been found that valuations relate to amygdala activations (see, e.g., Morrison and Salzman 2010; Murray 2007; Salzman and Fusi 2010). As valuing can be seen as a grounding for a decision, it turns out that it concerns an emotional grounding. Bad decisions are those that are not solidly grounded by having a positive feeling about them. They may not last long, as any opportunity to get rid of them will be a temptation to reconsider the decision. Recent neurological literature addressing this idea of emotional valuing and grounding of decisions relates the notion of value to the amygdala; e.g., Bechara et al. (2003), Bechara et al. (1999), Montague and Berns (2002), Janak and Tye (2015),

Jenison et al. (2011), Morrison and Salzman (2010), Ousdal et al. (2014), Pessoa (2010), Rangel et al. (2008) and Rudebeck and Murray (2014).

So, in this perspective emotions and the assumed rationality of decision making are assumed to go hand in hand. This theme is the focus of this chapter. The question addressed is how computational models for emotion generation can be used to model decision making processes, and how it can be analysed how rational such a decision making model is. So, it will be computationally addressed how emotions are not an enemy but a vehicle for rationality, and it will be analysed in what sense emotional valuing as a basis for decision making satisfies some rationality criterion.

To analyse this it has to be taken into account that decision making is not just an instantaneous process in the present, but it has an embedding in the temporal dimension. In this temporal dimension experiences with the environment in earlier situations are relevant, as they have implications for future decision situations: adaptive processes tune the decision making mechanism to such experiences. These experiences relate to emotions and by strengthening the associations to such emotions, in future decision situations the effect of these emotions will be stronger. As a consequence the decisions will be adapted better to the environment. By such an adaptive process the decisions in some way become more rational, given the increasing amount of knowledge about the environment built up by experiences over time. Sometimes it can be observed that within sport, athletes may show a strong emotion after they made an effort. For example, tennis players after they have hit a decisive ball: when successful a positive emotion occurs, and when not successful a negative emotion. Given the above analysis, by learning these emotions get a stronger association to the specific performance. This will already work when feeling the emotions without expressing them. But perhaps expressing the emotions will provide additional strengthening, as via the body loop (see Chap. 3, Sect. 3.2) this will increase the level of emotion felt, and therefore will strengthen the learning process.

In this chapter in particular it will be analyzed to which extent a biologically plausible emotion generation model together with a biologically plausible learning model is able to model decision making processes that satisfy some rationality criterion.

The adaptive temporal-causal network model for decision making considered in this chapter, first generates preparations for a number of options relevant for a situation at hand. By internal simulation the effects of these options are predicted as sensory representations. Then emotions are generated: these predicted effects trigger preparations for emotional responses and based on predictive as-if body loops, associated feeling states are generated. Thus emotional valuations of the options are obtained; e.g., Damasio (1994, 2004, 2010), Janak and Tye (2015), Jenison et al. (2011), Ousdal et al. (2014), Pearson et al. (2014), Pessoa (2010) and Rangel et al. (2008). The activation level of such a feeling state (which is considered as positive), strengthens the preparation for the related option. This process leads to an emerging strongest option as an outcome of the decision.

The type of biologically inspired learning modeled incorporated in a slightly simplified version of the decision model is a Hebbian learning model (e.g., Hebb 1949;

Gerstner and Kistler 2002), which has been applied to different types of connections in the decision model: not only to the association from stimulus to preparation (as in classical conditioning), but also to the connections in the as-if body loops that represent the strength of the association between preparation and feeling.

The next question addressed is how it can be evaluated to which extent a specific decision making model can be considered as being rational. To this end two different notions of rationality are defined, depending on characteristics of the environment. What is rational indeed depends strongly on the characteristics of the environment, as this environment determines what the effects of a chosen action will be. For example, buying fruit in some shop may lead to quite positive feelings if the quality of this fruit is good, but not when this quality is bad. Therefore a rationality measure has to relate to the environment's characteristics in the sense of the environment's behaviour when actions are performed. Two examples of such a rationality measure were defined and have been applied to evaluate the computational decision model. These notions of rationality are based on the assumption that the more the person makes choices that are the most beneficial within the given environment, the more rational it is.

In this chapter, in Sect. 6.2 the adaptive temporal-causal network model for decision making is introduced. Sections 6.3–6.5 present a number of simulation results for a deterministic world, a stochastic world and a changing world, respectively. In Sect. 6.6 measures for rationality are discussed, and the adaptive network model is evaluated with respect to these measures. Finally, Sect. 6.7 is a discussion.

6.2 The Adaptive Temporal-Causal Network Model for Decision Making

The adaptive temporal-causal network model described here is based on neurological notions such as valuing of decision options in relation to feeling, and internal simulation loops and execution loops for each of the options. These loops cover as-if body loops and body loops involving generated feelings.

Although an important function attributed to the amygdala concerns fear, in recent years also a function beyond this fear context has been found; e.g., Lindquist and Barrett (2012), Murray (2007) and Pessoa (2010). Parts of the prefrontal cortex (PFC) and other brain areas such as hippocampus, basal ganglia, and hypothalamus have extensive, often have bidirectional connections with the amygdala; see, for example, Ghashghaei et al. (2007), Janak and Tye (2015), Likhtik and Paz (2015), Morrison and Salzman (2010) and Salzman and Fusi (2010). Usually stimuli lead to associated responses and their predicted effects, which in turn trigger emotional responses relating to a rewarding or aversive feeling. These feelings represent a way of experiencing the value of such a predicted effect of a response: the extent to which it is positive or negative for the person. The underlying idea of neurologically represented value also plays a central role in neuroeconomics, for example, in work on the neural basis of economic choice. Much literature addresses the way in which

in decision-making different options are compared, and choices are related to a notion of value as represented in the amygdala; e.g., Bechara et al. (2003), Bechara et al. (1999), Montague and Berns (2002), Morrison and Salzman (2010), Ousdal et al. (2014), Pessoa (2010), Rangel et al. (2008) and Sugrue et al. (2005). For more details about emotions and their role on other mental processes, also see Chap. 3.

In Fig. 6.1 it is shown how a sensed stimulus s leads to its sensory representation which in turn triggers a preparation state for one or more actions a . Note that the adaptive element has been left out of consideration here. Such a preparation state leads to a sensory representation of a predicted effect e : internal simulation of action a . By this sensory representation of e , preparations for emotional responses b are triggered for which a prediction is made, thus generating a feeling state b : as-if body loop for internal simulation of body state b . Feeling these emotions represents a way of experiencing the value of such a prediction: to which extent it is positive or negative. This feeling state serves as a valuation, and in turn amplifies the activation of the preparation for the concerning option. The pattern involves two cycles, as depicted in Fig. 6.1. Table 6.1 shows a conceptual matrix representation for this model.

Note that these cycles are active in parallel for all options a (partially) triggered by s , and result in specific valuation values for each of the options, which, depending on their value, strengthen the options. In this way the valuing process strengthens some of these option preparations more due to higher values of the connections involved in the two cycles, such as the connection from valuation to preparation, or the connections involved in prediction and in the as-if body loop.

Note that this pattern is similar to what is considered in Damasio (1994)'s Somatic Marker Hypothesis. Moreover, note that this provides a (circular) causal role of feelings in decisions for actions, which has some parallel to the discussion in Philosophy of Mind on the causal role of qualia; e.g., Kim (1996).

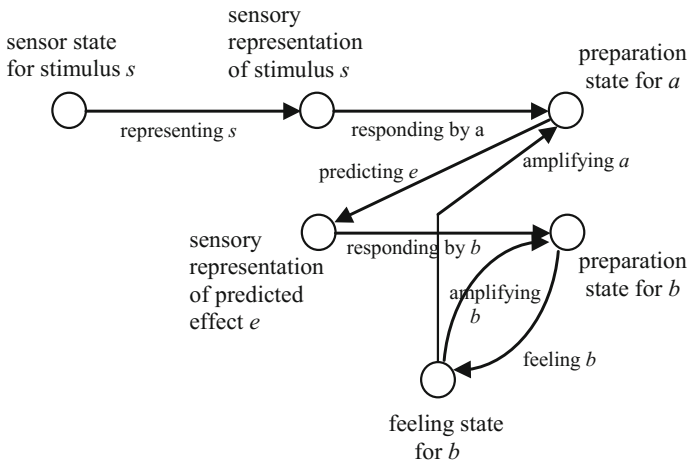


Fig. 6.1 Conceptual graphical representation of a temporal-causal network model for decision making based on emotion-related valuing

Table 6.1 Conceptual matrix representation of a temporal-causal network model for decision making based on emotion-related valuing

To	ss_s	srs_s	ps_a	srs_e	ps_b	fs_b
From						
ss_s		$\omega_{representing}$				
srs_s			$\omega_{responding_a}$			
ps_b				$\omega_{predicting}$		
srs_e					$\omega_{responding_b}$	
ps_b						$\omega_{feeling}$
fs_b			$\omega_{amplifying_a}$		$\omega_{amplifying_b}$	
η_Y	–	η_{srs_s}	η_{ps_b}	η_{srs_e}	η_{ps_b}	η_{fs_b}
$c_Y(\dots)$	–	$c_{srs_s}(V)$	$c_{ps_a}(V_1, V_2)$	$c_{srs_e}(V)$	$c_{ps_b}(V_1, V_2)$	$c_{fs_b}(V)$

The numerical model is shown in Box 6.1, with connection weights $\omega_{X,Y} \geq 0$. Note that the connection weights determine the valuation generated for a given option a . For example, when in a special case for a given option a all of $\omega_{responding_a}$, $\omega_{predicting}$, $\omega_{responding_b}$, $\omega_{feeling}$, and $\omega_{amplifying_a}$ are higher than those for all other options, then this option a will get the highest activation level for feeling fs_b and for ps_a . On the other hand, it may well be the case that some option a has the highest $\omega_{responding_a}$, but still gets the lowest valuation by the feeling level of fs_b and as a consequence lowest activation level ps_a , due to lower values for other connections involved in the cycles for this option. In such a case, although at forehand there are good indications for a as a response for the given stimulus s , still because of negative valuation of the predicted effect it is not pursued.

LP1 Representing stimulus s

$$dsrs_s/dt = \eta_{srs_s} [c_{srs_s}(\omega_{representing} ss_s) - srs_s]$$

$$srs_s(t + \Delta t) = srs_s(t) + \eta_{srs_s} [c_{srs_s}(\omega_{representing} ss_s(t)) - srs_s(t)] \Delta t$$

LP2 Generating and amplifying response a

$$dps_a/dt = \eta_{ps_a} [c_{ps_a}(\omega_{responding_a} srs_s, \omega_{amplifying_a} fs_b) - ps_a]$$

$$ps_a(t + \Delta t) = ps_a(t) + \eta_{ps_a} [c_{ps_a}(\omega_{responding_a} srs_s(t), \omega_{amplifying_a} fs_b(t)) - ps_a(t)] \Delta t$$

LP3 Generating and amplifying response b

$$dps_b/dt = \eta_{ps_b} [c_{ps_b}(\omega_{responding_b} srs_e, \omega_{amplifying_b} fs_b) - ps_b]$$

$$ps_b(t + \Delta t) = ps_b(t)$$

$$+ \eta_{ps_b} [c_{ps_b}(\omega_{responding_b} srs_e(t), \omega_{amplifying_b} fs_b(t)) - ps_b(t)] \Delta t$$

LP4 Predicting effect e

$$ds_{rs_e}/dt = \eta_{s_{rs_e}} [c_{s_{rs_e}}(\omega_{\text{predicting}} ps_a) - s_{rs_e}]$$

$$s_{rs_e}(t + \Delta t) = s_{rs_e}(t) + \eta_{s_{rs_e}} [c_{s_{rs_e}}(\omega_{\text{predicting}} ps_a(t)) - s_{rs_e}(t)] \Delta t$$

LP5 Feeling b

$$dfs_b/dt = \eta_{fs_b} [c_{fs_b}(\omega_{\text{feeling}} ps_b) - fs_b]$$

$$fs_b(t + \Delta t) = fs_b(t) + \eta_{fs_b} [c_{fs_b}(\omega_{\text{feeling}} ps_b(t)) - fs_b(t)] \Delta t$$

The symbols are explained as follows:

- ss_s sensor state for stimulus s
- srs_s sensory representation state for stimulus s
- ps_a preparation state for decision option a
- ps_b preparation state for emotional response b
- srs_e sensory representation state for predicted effect e
- fs_b feeling state for b

Box 6.1 Numerical representation of a temporal-causal network model for decision making based on emotion-related valuing

The analysis of rationality of decision models based on emotional valuing has been based on a simplified variant of the decision making model discussed above. This simplified variant was adopted from Treur and Umair (2011, 2015). For this variant it has been analysed how adaptivity can be added and in how far this makes the model behave rationally for a given environment.

A graphical conceptual representation of the model of this adaptive temporal-causal network model is depicted in Fig. 6.2. This picture also shows formal name representations explained in Table 6.2. The numerical representation of the model is given in the detailed specifications below. A main difference with the model depicted in Fig. 6.1 is that the causal chain

sensory representation of predicted effect $e \rightarrow$ preparation state for emotion $b \rightarrow$ feeling b

is shortened into a causal chain

sensory representation of predicted effect $e \rightarrow$ feeling e

Here ‘feeling e ’ plays the role of the emotion effect b related to e . Moreover, a subscript i is added to represent different options a_i for $i = 1, 2, \dots$. Furthermore, the weights of the connections to which Hebbian learning is applied are indicated by $\omega_{1,i}$ (from stimulus to response), $\omega_{2,i}$ (from feeling to preparation), and $\omega_{3,i}$

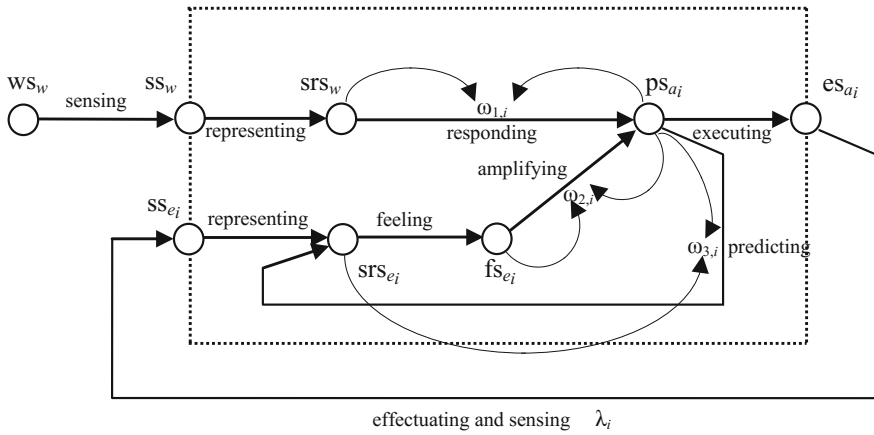


Fig. 6.2 Graphical conceptual representation of the adaptive temporal-causal network model for decision making, which will be evaluated from a rationality perspective

Table 6.2 Overview of the (dynamic) states and adaptive connections used in the model

Formal name	Informal name	Description
ws_w	World state for w	This characterizes the current external world situation the person is facing
ss_w	Sensor state for w	The person observes the world state through the sensor state, which provides sensory input
ss_{e_i}	Sensor state for effect e_i	The person observes the effect e_i through the sensor state, which provides sensory input
srs_w	Sensory representation state for w	Internal representation of w
srs_{e_i}	Sensory representation state for effect e_i	A feeling state fs_{e_i} is generated, via the predicted sensory representation srs_{e_i} for action a_i . This provides a valuing for action option a_i
fs_{e_i}	Feeling state for effect e_i	
ps_{a_i}	Preparation state for action a_i	Preparation for a response a_i
es_{a_i}	Execution state for a_i	Execution of a_i
$\omega_{1,i}$	Weight of the responding connection from srs_w to ps_{a_i}	This models the relation between stimulus w and directly associated <i>response</i> a_i
$\omega_{2,i}$	Weight of the amplifying connection from fs_{e_i} to ps_{a_i}	This models how the generated feeling affects (<i>amplifies</i>) the preparation for response a_i
$\omega_{3,i}$	Weight of the predicting connection from ps_{a_i} to srs_{e_i}	This models how the preparation for response a_i affects (<i>predicts</i>) the representation for effect e_i

(prediction of the effect). Note that for three options a_i for $i = 1, 2, 3$, these are 9 connections that can be learnt.

To get an idea of a real life context for the model, consider a simple scenario in which a person can choose among three options a_i with $i = 1, 2$, or 3, indicating

going to three different shops to get a certain product, for example, fruit. In the world each of the three options provides this product with a certain satisfaction factor, indicated by the world's characteristics or *effectiveness rates* λ_i with $i = 1, 2$, or 3 (λ_i measured between 0 and 1) for the three options respectively. The model describes a situation in which a number of times the person goes to all three shops to find out (and learn) how these satisfaction factors are. In each of the shops the person buys a fraction of the fruit corresponding to the person's tendency or preference to decide for this shop.

For each i the execution state for option a_i combined with the (possibly stochastic) effectiveness of executing a_i in the world (indicated by λ_i) activates the sensor state for effect e_i via a prediction link as described above (internal simulation). Each of the preparations for a_i generates a level of feeling for e_i . This functions as a valuation of the prediction of the action effect by the internal simulation. Therefore in turn it affects the level of the related preparation for a_i , which creates a recursive loop. Dynamic interaction within each of these three loops results in an equilibrium state for the strength of the preparation and of the feeling. Depending on these values, the option a_i is actually activated with a certain intensity. The specific strengths of the connections from the sensory representation to the preparations, and within the recursive loops can be innate, or are acquired in an adaptive manner during lifetime by learning. The adaptivity in the model is based on Hebbian learning.

The numerical representation of the model is presented below through local (dynamic) properties LP0 to LP9. As a first step sensing of the world state w takes place (see also Fig. 6.2, the arrow in the left upper corner, with label sensing). This means that the activation value of the sensor state for w is updated (gradually) to get its value more in the direction of the value of the world state for w (the weight of the connection is set on 1). After a few steps these activation values will become (practically) equal. This is expressed in property LP0.

LP0 Sensing a world state

$$\begin{aligned} dss_w(t)/dt &= \eta_0[ws_w(t) - ss_w(t)] \\ ss_w(t + \Delta t) &= ss_w(t) + \eta_0[ws_w(t) - ss_w(t)]\Delta t \end{aligned}$$

From the sensor state for w , the sensory representation of w is updated by dynamic property LP1. Again this means that the activation value of the sensory representation for w is updated (gradually) to get its value more in the direction of the value of the sensor state of w (also this link gets weight 1). After a few steps these values will become (practically) equal. This is expressed in property LP1.

LP1 Representing a sensed world state

$$\begin{aligned} dsrs_w(t)/dt &= \eta_0[ss_w(t) - srs_w(t)] \\ srs_w(t + \Delta t) &= srs_w(t) + \eta_0[ss_w(t) - srs_w(t)]\Delta t \end{aligned}$$

To get an updated value of the preparation state for action a_i , there are impacts of two other states (see the two arrows pointing at the preparation state in Fig. 6.2, right upper corner): the sensory representation of w and the feeling state for effect e_i . This means that a combination of the activation values of these states has impact on the activation value of the preparation state. The combination function $c(\cdot)$ to aggregate the two impacts is the advanced logistic function **alogistic** $_{\sigma,\tau}(V_1, V_2)$, thus keeping the resultant value in the range $[0, 1]$:

$$\mathbf{alogistic}_{\sigma,\tau}(V_1, V_2) = \left(\frac{1}{1 + e^{-\sigma(V_1 + V_2 - \tau)}} - \frac{1}{1 + e^{\sigma\tau}} \right) (1 + e^{-\sigma\tau})$$

where σ is the steepness and τ is the threshold value.

Using this, dynamic property LP2 describes the update of the preparation state for a_i upon impact from the sensory representation of w and feeling state for effect e_i .

LP2 Generating and amplifying a response a_i

$$\begin{aligned} \mathbf{dps}_{a_i}(t)/\mathbf{dt} &= \eta_1 [\mathbf{alogistic}_{\sigma_i,\tau_i}(\omega_{1,i}(t)\mathbf{srs}_w(t), \omega_{2,i}(t) \mathbf{fs}_{e_i}(t)) - \mathbf{ps}_{a_i}(t)] \\ \mathbf{ps}_{a_i}(t + \Delta t) &= \mathbf{ps}_{a_i}(t) + \eta_1 [\mathbf{alogistic}_{\sigma_i,\tau_i}(\omega_{1,i}(t)\mathbf{srs}_w(t), \omega_{2,i}(t) \mathbf{fs}_{e_i}(t)) - \mathbf{ps}_{a_i}(t)]\Delta t \end{aligned}$$

In a similar manner dynamic property LP3 describes the update of the activation level of the sensory representation of effect e_i upon impact from the values for the preparation state for a_i and the sensor state for e_i (see Fig. 6.2, left under corner).

LP3 Representing effect e_i based on predicted and sensed effect

$$\begin{aligned} \mathbf{dsrs}_{e_i}(t)/\mathbf{dt} &= \eta_2 [\mathbf{alogistic}_{\sigma_i,\tau_i}(\omega_{3,i}\mathbf{ps}_{a_i}(t), \mathbf{ss}_{e_i}(t)) - \mathbf{srs}_{e_i}(t)] \\ \mathbf{srs}_{e_i}(t + \Delta t) &= \mathbf{srs}_{e_i}(t) + \eta_2 [\mathbf{alogistic}_{\sigma_i,\tau_i}(\omega_{3,i}\mathbf{ps}_{a_i}(t), \mathbf{ss}_{e_i}(t)) - \mathbf{srs}_{e_i}(t)]\Delta t \end{aligned}$$

Dynamic property LP4 describes the update of the activation level of the feeling state for effect e_i upon impact from the activation level of the sensory representation for e_i (see Fig. 6.2 in the middle under).

LP4 Feeling for effect e_i

$$\begin{aligned} \mathbf{dfs}_{e_i}(t)/\mathbf{dt} &= \eta_0 [\mathbf{srs}_{e_i}(t) - \mathbf{fs}_{e_i}(t)] \\ \mathbf{fs}_{e_i}(t + \Delta t) &= \mathbf{fs}_{e_i}(t) + \eta_0 [\mathbf{srs}_{e_i}(t) - \mathbf{fs}_{e_i}(t)]\Delta t \end{aligned}$$

LP5 describes how the activation level of the execution state for action a_i is updated upon impact from the activation level of its preparation state (see Fig. 6.2, right upper corner).

LP5 Executing action a_i

$$\begin{aligned} \mathbf{des}_{a_i}(t)/\mathbf{dt} &= \eta_0 [\mathbf{ps}_{a_i}(t) - \mathbf{es}_{a_i}(t)] \\ \mathbf{es}_{a_i}(t + \Delta t) &= \mathbf{es}_{a_i}(t) + \eta_0 [\mathbf{ps}_{a_i}(t) - \mathbf{es}_{a_i}(t)]\Delta t \end{aligned}$$

LP6 describes the update of the activation level of the sensor state for effect e_i upon impact from the activation level of the execution state for action a_i (see Fig. 6.2, the arrow from right upper corner to left under corner). Here the world characteristics λ_i play their role.

LP6 Effectuating and sensing effect e_i of action a_i

$$\begin{aligned} dss_{e_i}(t)/dt &= \eta_0[\lambda_i es_{a_i}(t) - ss_{e_i}(t)] \\ ss_{e_i}(t + \Delta t) &= ss_{e_i}(t) + \eta_0[\lambda_i es_{a_i}(t) - ss_{e_i}(t)]\Delta t \end{aligned}$$

Just to get the idea, Fig. 6.3 shows some of the variables for a simple instance of the scenario and the output of the basic model in a deterministic environment without learning. This Fig. 6.3 only shows the output of the execution states for the options a_i over time without learning any of the connections, for given deterministic world characteristics or preferences for the available options. The horizontal axis represents time (corresponds to simulation steps). The vertical axis represents the activation values of the depicted states which can be any values between 0 and 1. The upper left hand graph indicates the recurring world state ws_w for which the decision problem and the options occur. This serves as input. The activation value for this world state switches from 0 to 1 and back only. The upper right hand graph shows the (different) world characteristics λ_i for the three options. They are preset as world characteristics for any given scenario. The lower left hand graph shows that the strengths ω_{1i} for all three connections from srs_w to ps_{a_i} are the same for this case and constant for the three options (no learning takes place). The lower right hand graph shows that the tendencies to choose for the three responses or decision options a_1 , a_2 or a_3 for w remain the same (due to the fact that no learning takes place).

For the considered case study it was assumed that three options are available to the person and the objective is to see how rationally a person makes his or her decisions using the model as described here, under deterministic as well as stochastic world characteristics and for static as well as changing worlds. The dynamic properties LP7 to LP9 describe a learning mechanism for three connection weights:

- (A) connection weight ω_{1i} for the *responding* connection from sensory representation state for w to preparation state for action option a_i
- (B) connection weight ω_{2i} for the *amplifying* connection from feeling state for effect e_i to preparation state for action option a_i
- (C) connection weight ω_{3i} for the *predicting* connection from preparation state for a_i to sensory representation state for (predicted) effect e_i

These have been explored separately (A), (B), or (C), and in combination (ABC). The first type of learning (A) corresponds to learning as considered traditionally for an association between stimulus and response. The other two types of learning (B and C) concern strengthening of the predictive and valuing loop. This models the association of the feeling for effect e_i to option a_i and how this feeling has impact on the preparation state for a_i as a way of valuing for this option.

The learning model used is Hebbian learning (Hebb 1949; Gerstner and Kistler 2002). This is based on the principle that strengthening of a connection between neurons over time may take place when both nodes are often active simultaneously ('neurons that fire together wire together'); see also Chap. 2, Sect. 2.10. More specifically, here it is assumed that the strength ω of a connection between states X_1 and X_2 is adapted using the following *Hebbian learning rule*, taking into account a maximal connection weight 1, a *learning rate* $\eta > 0$, and an *extinction rate* $\zeta \geq 0$ (usually small), and activation levels $X_1(t)$ and $X_2(t)$ of the two states involved. The first expression is in differential equation format, the last one in difference equation format

$$\begin{aligned} d\omega(t)/dt &= \eta X_1(t)X_2(t)(1 - \omega(t)) - \zeta\omega(t) \\ \omega(t + \Delta t) &= \omega(t) + [\eta X_1(t)X_2(t)(1 - \omega(t)) - \zeta\omega(t)]\Delta t \end{aligned}$$

Such Hebbian learning rules can be found, for example, in Gerstner and Kistler (2002, p. 406). By the factor $1 - \omega$ the learning rule keeps the level of ω bounded by 1. When the extinction rate ζ is relatively low, the upward changes during learning are proportional to both $X_1(t)$ and $X_2(t)$ and maximal learning takes place when both are 1. Whenever one of these activation levels is 0 (or close to 0) extinction takes over, and ω slowly decreases (unlearning).

First, based on the Hebbian learning mechanism described above local dynamic property LP7 models the update of the strength of the connection from sensory representation of world state w to preparation of a_i (type A; see Fig. 6.2 the arrow with labels responding and $\omega_{1,i}$ in the upper part, in the middle).

LP7 (A) Learning the responding links from representation of w to preparation for a_i

$$\begin{aligned} d\omega_{1,i}(t)/dt &= \eta srs_w(t)ps_{a_i}(t)(1 - \omega_{1,i}(t)) - \zeta\omega_{1,i}(t) \\ \omega_{1,i}(t + \Delta t) &= \omega_{1,i}(t) + [\eta srs_w(t)ps_{a_i}(t)(1 - \omega_{1,i}(t)) - \zeta\omega_{1,i}(t)]\Delta t \end{aligned}$$

Similarly the learning of the two connections involved in the predicting and valuing loop (B and C) are specified in LP8 and LP9 (see Fig. 6.2, the arrows with label $\omega_{2,i}$ (B) and $\omega_{3,i}$ (C)).

LP8 (B) Learning the amplifying links from feeling for e_i to preparation for a_i

$$\begin{aligned} d\omega_{2,i}(t)/dt &= \eta fs_{e_i}(t)ps_{a_i}(t)(1 - \omega_{2,i}(t)) - \zeta\omega_{2,i}(t) \\ \omega_{2,i}(t + \Delta t) &= \omega_{2,i}(t) + [\eta fs_{e_i}(t)ps_{a_i}(t)(1 - \omega_{2,i}(t)) - \zeta\omega_{2,i}(t)]\Delta t \end{aligned}$$

LP9 (C) Learning the predicting links from preparation of a_i to representation of e_i

$$\begin{aligned} d\omega_{3,i}(t)/dt &= \eta ps_{a_i}(t)srs_{e_i}(t)(1 - \omega_{3,i}(t)) - \zeta\omega_{3,i}(t) \\ \omega_{3,i}(t + \Delta t) &= \omega_{3,i}(t) + [\eta ps_{a_i}(t)srs_{e_i}(t)(1 - \omega_{3,i}(t)) - \zeta\omega_{3,i}(t)]\Delta t \end{aligned}$$

6.3 Simulation Results for a Deterministic World

In this section some of the simulation results, performed using numerical software, are described in detail. A real life scenario context that can be kept in mind to understand what is happening is the following.

Scenario context

At regular times (for example, every week) a person faces the decision problem where to buy a certain type of fruit. There are three options for shops to choose from. Each option i has its own characteristics (price and quality, for example); based on these characteristics buying in this shop provides a satisfaction factor λ_i between 0 and 1. For the deterministic case it is just this value that is obtained for this option. For the stochastic case a probability distribution around this λ_i will determine the outcome. For a static world the λ_i remain constant, whereas for a dynamic world they may change over time. By going to a shop the person experiences the related outcome λ_i and learns from this. To find out which shop is most satisfactory, each time the person decides to go to all three of them to buy some fraction of the amount of fruit needed. This fraction corresponds to the person's tendency or preference to decide for this shop. Due to the learning this tendency changes over time, in favour of the shop(s) which provide(s) the highest satisfaction.

The simulation results address different scenarios reflecting different types of world characteristics, from worlds that are deterministic to stochastic, and from a static to a changing world. Moreover, learning the connections was done one at a time (A), (B), (C), and learning multiple connections simultaneously (ABC). An overall summary of the results is given in Table 6.2. Note that this table only contains the values of the connection weights and activation levels of the execution states after completion of the simulation experiments: at the end time. In contrast in the current and next section the processes are shown performed to achieve these final values.

Results for the rationality measures are presented in the next section. For all simulation results shown, time is on the horizontal axis whereas the vertical axis shows the activation level of the different states. Step size for all simulations is $\Delta t = 1$. Figure 6.3 shows simulation results for the model under constant, deterministic world characteristics: $\lambda_1 = 0.9$, $\lambda_2 = 0.2$, and $\lambda_3 = 0.1$. Other parameters are set as: learning rate $\eta = 0.04$, extinction rate $\zeta = 0.0015$, initial connection weight $\omega_{2i} = \omega_{3i} = 0.8$, speed factors $\eta_0 = 1$, $\eta_1 = 0.5$, $\eta_2 = 1$, steepness $\sigma = 2$ and threshold $\tau = 1.2$ for preparation state, and $\sigma = 10$ and $\tau = 0.3$ for sensory representation of e_i . For the initial 80 time units the stimulus w is kept 1 and for next 170 time units it is kept 0 and the same sequence of activation and deactivation for the stimulus is repeated for the rest of the simulation (Fig. 6.4).

Moreover, it depicts the situation in which only one type of link with weight ω_{1i} is learned as specified in LP7 using the Hebbian approach (A) for the *responding* connections from sensory representation of w to the preparation states for the options a_i . It is shown that the model adapts the connection weights ω_{1i} of these links according to the world characteristics given by λ_i . So ω_{11} strengthens more and more over time, resulting in the higher activation level of the execution state for

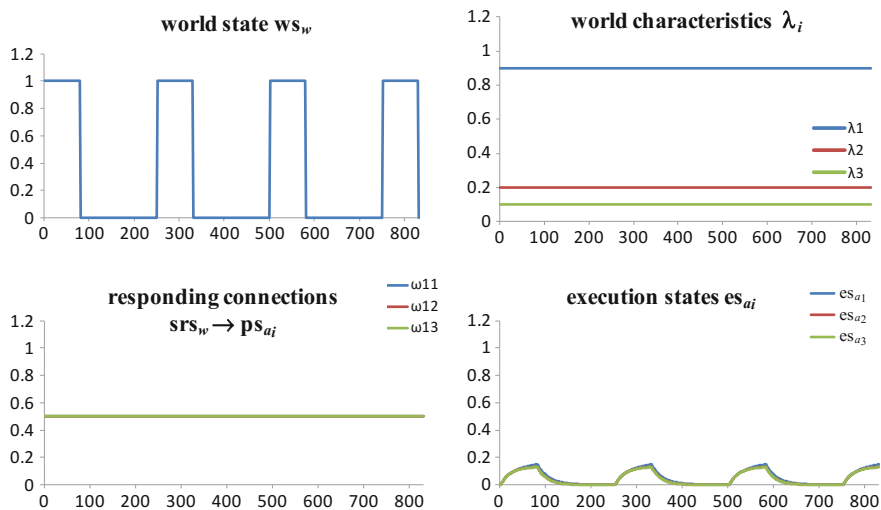


Fig. 6.3 Output of the basic model without learning

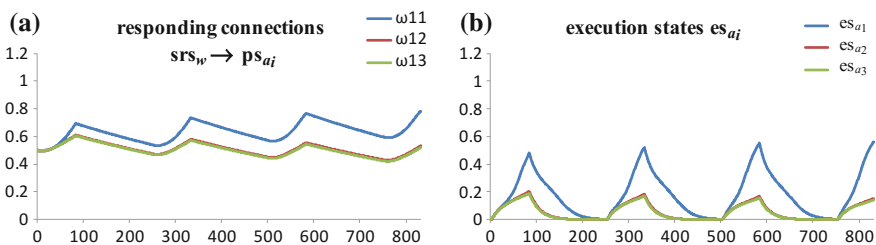


Fig. 6.4 Deterministic world: **a** responding connection weights (A) **b** execution states for a_i . Initial values $\omega_{1,1} = \omega_{1,2} = \omega_{1,3} = 0.5$; $\eta = 0.04$, $\zeta = 0.0015$

a_1 compared to the activation level of the execution states for the other two options a_2 and a_3 .

Figures 6.5 and 6.6 show the simulation results while learning is performed for the weights of the *amplifying* links (B) from feeling to preparation state for a_i and the *predicting* links (C) from preparation state to sensory representation of e_i , respectively.

Figure 6.7 shows the results when the Hebbian learning is applied on all links simultaneously (ABC). These results show that for deterministic world characteristics the person successfully adapts the connection weight in all four different cases rationally, as these connection weights become more in line with the world characteristics.

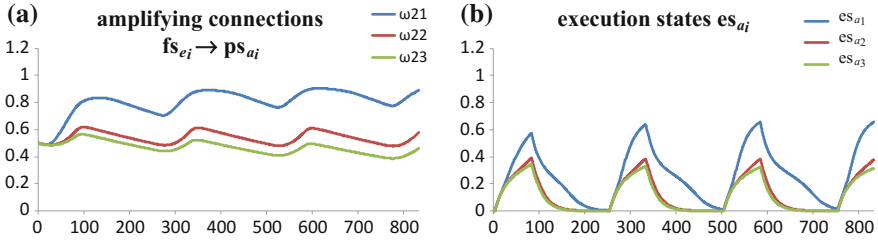


Fig. 6.5 Deterministic world: **a** amplifying connection weights (B) **b** execution states for a_i . Initial values $\omega_{2,1} = \omega_{2,2} = \omega_{2,3} = 0.5$, $\eta = 0.04$; $\zeta = 0.0015$

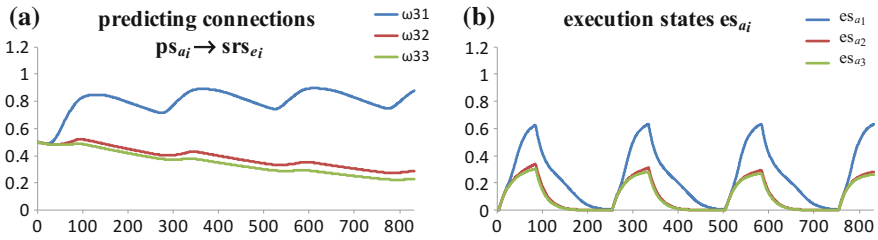


Fig. 6.6 Deterministic world: **a** predicting connection weights (C) **b** execution states for a_i . Initial values $\omega_{3,1} = \omega_{3,2} = \omega_{3,3} = 0.5$, $\eta = 0.04$, $\zeta = 0.0015$

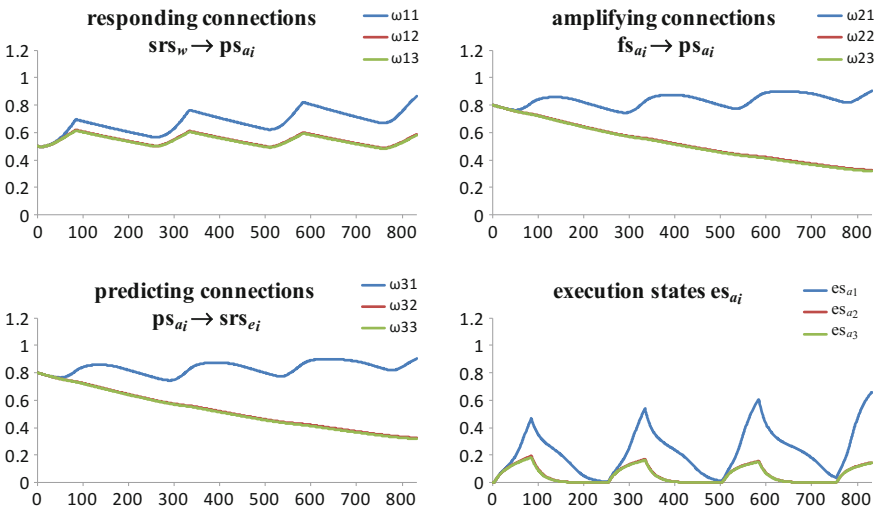


Fig. 6.7 Deterministic world: weights for all three connections (ABC) and execution states for a_i . Initial values $\omega_{1,1} = \omega_{1,2} = \omega_{1,3} = 0.5$, $\omega_{2,1} = \omega_{2,2} = \omega_{2,3} = 0.8$, $\omega_{3,1} = \omega_{3,2} = \omega_{3,3} = 0.8$; $\eta_i = 0.04$, $\zeta_i = 0.0013$

6.4 Simulation Results for a Stochastic World

Other experiments were carried out for a stochastic world with four different cases as mentioned earlier. To simulate the stochastic world, probability distribution functions (PDF) were defined for λ_i according to a Normal Distribution. Using these PDFs, the random numbers were generated for λ_i limiting the values for the interval $[0, 1]$ with $\mu_1 = 0.9$, $\mu_2 = 0.2$ and $\mu_3 = 0.1$ for the λ_i respectively. Furthermore the standard deviation for all λ_i was taken 0.1. Figure 6.8 shows the world state w and stochastic world characteristics λ_i . Figures 6.9, 6.10, 6.11 and 6.12 show the simulation results while learning is performed for the links (A) from sensory representation of w to preparation state for a_i , (B) from feeling to

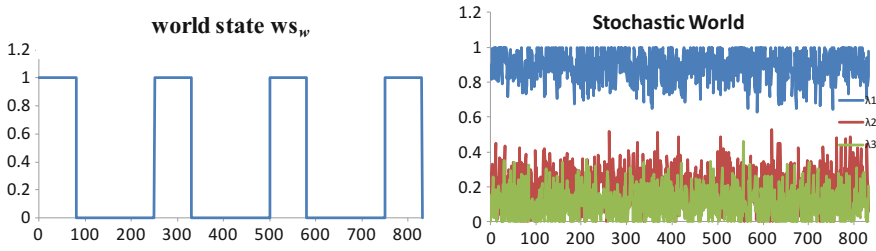


Fig. 6.8 Stochastic world

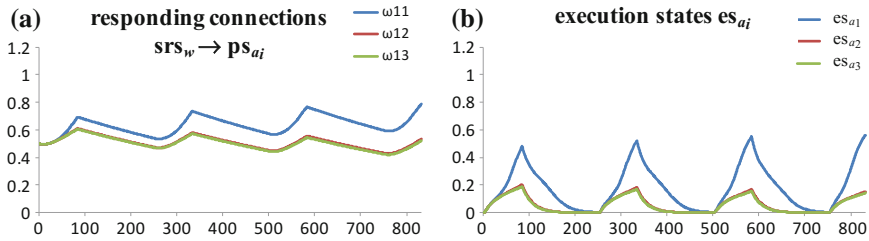


Fig. 6.9 Stochastic world: **a** responding connection weights (A) **b** execution states. Initial values $\omega_{1,1} = \omega_{1,2} = \omega_{1,3} = 0.5$; $\eta = 0.04$, $\zeta = 0.0015$

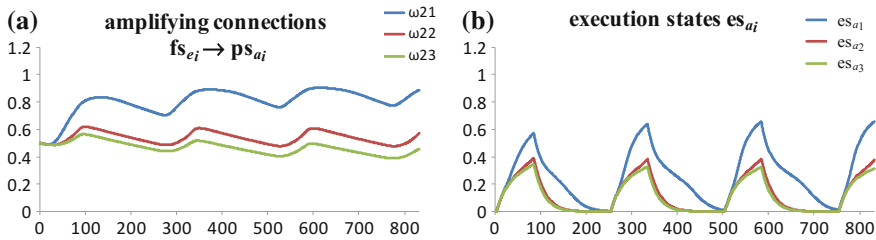


Fig. 6.10 Stochastic world: **a** amplifying connection weights (B) **b** execution states. Initial values $\omega_{2,1} = \omega_{2,2} = \omega_{2,3} = 0.5$; $\eta = 0.04$, $\zeta = 0.0015$

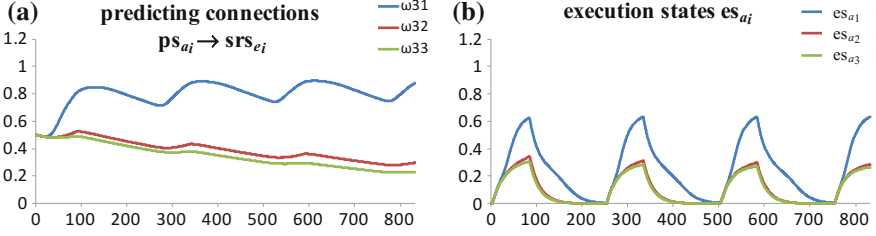


Fig. 6.11 Stochastic world: **a** predicting connection weights (C) **b** execution states. Initial values $\omega_{3,1} = \omega_{3,2} = \omega_{3,3} = 0.5$; $\eta = 0.04, \zeta = 0.0015$

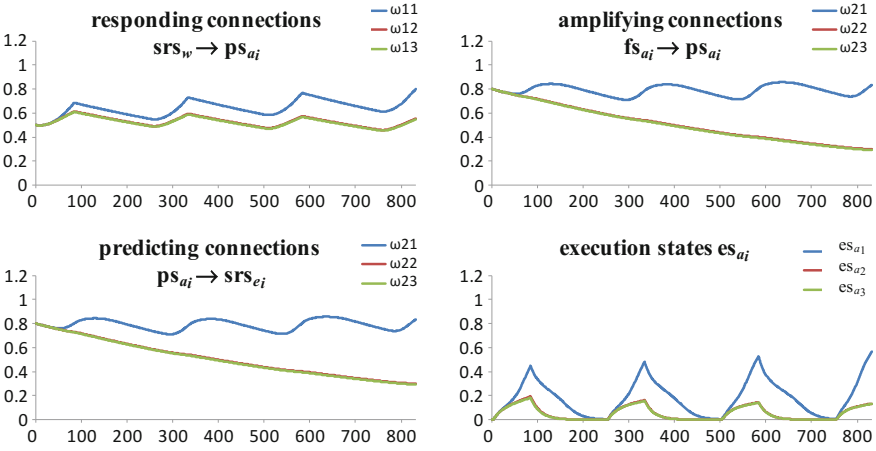


Fig. 6.12 Stochastic world: weights for all three connections (ABC) and execution states. Initial values $\omega_{1,1} = \omega_{1,2} = \omega_{1,3} = 0.5$, $\omega_{2,1} = \omega_{2,2} = \omega_{2,3} = 0.8$, $\omega_{3,1} = \omega_{3,2} = \omega_{3,3} = 0.8$; $\eta_i = 0.04$, $\zeta_i = 0.0013$

preparation state for a_i , and (C) from preparation state to sensory representation of e_i respectively, one at a time, and (ABC) all three.

It can be seen from these results that also in a stochastic scenario the temporal-causal network model successfully learnt the connections and adapted the connections and execution states to the world characteristics with results quite similar to the results for a deterministic world.

6.5 Simulation Results for a Changing Stochastic World

Another scenario was explored in which the (stochastic) world characteristics were changing drastically from $\mu_1 = 0.9$, $\mu_2 = 0.2$ and $\mu_3 = 0.1$ for the λ_i respectively to $\mu_1 = 0.1$, $\mu_2 = 0.2$ and $\mu_3 = 0.9$ for the λ_i respectively with standard deviation of 0.1 for all. Figures 6.13, 6.14, 6.15 and 6.16 show the results for such a scenario. The results show that the person has successfully adapted to the changing world

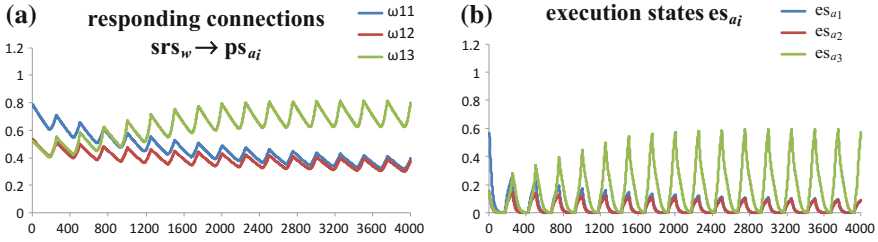


Fig. 6.13 Changing stochastic world: **a** responding connection weights (A) **b** execution states. Initial values $\omega_{1,1} = 0.78$, $\omega_{1,2} = 0.53$, $\omega_{1,3} = 0.52$; $\eta = 0.04$, $\zeta = 0.0015$

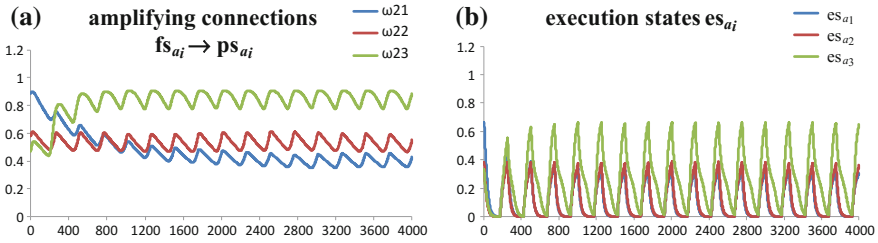


Fig. 6.14 Changing stochastic world: **a** amplifying connection weights (B) **b** execution states. Initial values $\omega_{2,1} = 0.88$, $\omega_{2,2} = 0.58$, $\omega_{2,3} = 0.47$; $\eta = 0.04$, $\zeta = 0.0015$

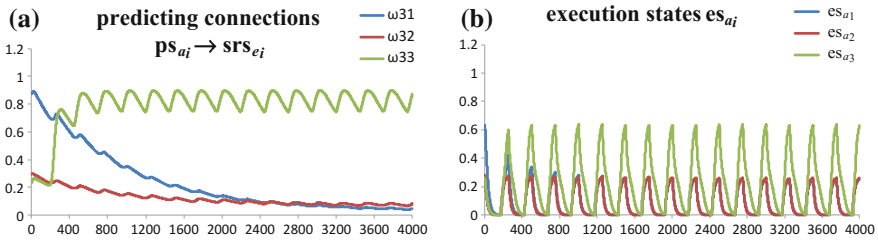


Fig. 6.15 Changing stochastic world: **a** predicting connection weights (C) **b** execution states. Initial values $\omega_{3,1} = 0.87$, $\omega_{3,2} = 0.30$, $\omega_{3,3} = 0.23$; $\eta = 0.04$, $\zeta = 0.0015$

characteristics over time, as the final execution states reflect more the changed world characteristics. The initial settings in this experiment were taken from the previous simulation results shown in Figs. 6.8 and 6.9 to keep the continuity of the experiment. It can be observed that the connection weight for option 3 becomes higher compared to the other options, and consequently the value of the execution state for a_3 becomes higher than for the other two by the end of experiment.

Note that the Table 6.3 contains the values of different connection weights and the activation level of execution states after the completion of simulation experiments (at the end time).

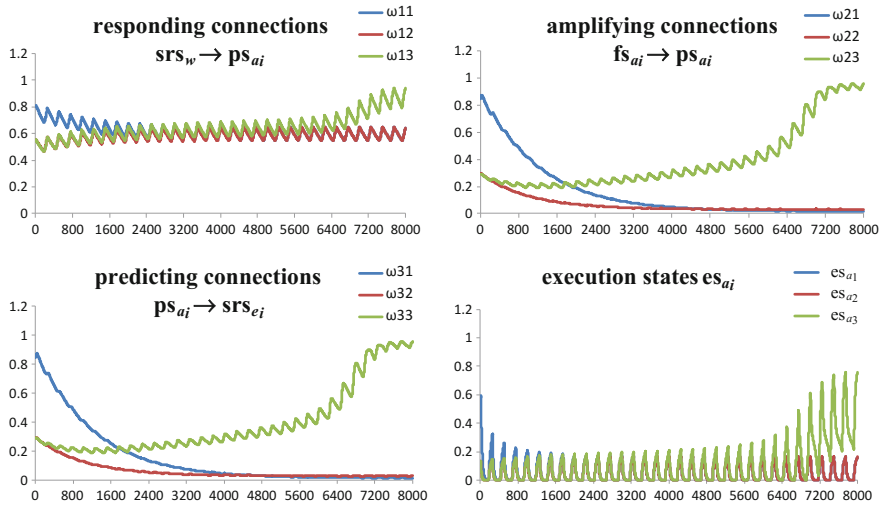


Fig. 6.16 Changing stochastic world: weights for all three connections (ABC) and execution states. Initial values $\omega_{1,1} = 0.80$, $\omega_{1,2} = 0.55$, $\omega_{1,3} = 0.54$, $\omega_{2,1} = \omega_{3,1} = 0.84$, $\omega_{2,2} = \omega_{3,2} = 0.30$, $\omega_{2,3} = \omega_{3,3} = 0.29$; $\eta_i = 0.04$, $\zeta_i = 0.001$

Table 6.3 Overview of the simulation results for all cases (A), (B), (C) and (ABC)

Connection	Scenario	ω_{x1}	ω_{x2}	ω_{x3}	ES ₁	ES ₂	ES ₃
Responding connection (A)	Deterministic	0.78	0.53	0.52	0.56	0.15	0.14
	Stochastic	0.78	0.53	0.52	0.56	0.15	0.14
	Changing	0.40	0.38	0.80	0.09	0.09	0.58
Amplifying connection (B)	Deterministic	0.89	0.58	0.46	0.65	0.38	0.31
	Stochastic	0.89	0.59	0.47	0.65	0.39	0.32
	Changing	0.42	0.57	0.89	0.30	0.37	0.65
Predicting connection (C)	Deterministic	0.88	0.29	0.23	0.63	0.28	0.26
	Stochastic	0.88	0.29	0.23	0.63	0.28	0.27
	Changing	0.04	0.08	0.87	0.25	0.26	0.63
All three connections (ABC)	Deterministic	0.81	0.55	0.54	0.59	0.13	0.13
		0.85	0.30	0.29			
		0.85	0.30	0.29			
	Stochastic	0.80	0.55	0.54	0.57	0.13	0.13
		0.84	0.30	0.29			
		0.84	0.30	0.29			
	Changing	0.64	0.64	0.94	0.16	0.16	0.75
		0.02	0.03	0.96			
		0.02	0.03	0.96			

6.6 Evaluating the Adaptive Temporal-Causal Network Model on Rationality

In the previous section it was shown that the temporal-causal network model behaves rationally in different scenarios. These scenarios and its different cases are elaborated in detail in the previous section, but the results were assessed with respect to their rationality in a qualitative and rather informal manner. For example, no attempt was made to assign an extent or level to the rationality observed during these experiments. The current section addresses this and to this end two different formally defined measures to assess the extent of the rationality are introduced. The notion of rationality aimed at in this chapter concerns that the person makes the choices that are most beneficial to it in the given environment. For example, it does not take into account the extent to which the person has information about the environment. So, if the person has had no experiences yet with the given environment (extreme form of incomplete information), then according to the notion considered here it will behave totally irrational. Only when the person gathers more information about the environment by having experiences with choices previously made, its behaviour will become more and more rational as an adaptation to its environment. The focus then is on how rational the person will become over time, in such an adaptation process to the environment. From the rationality measures considered here one rationality measure is based on a discrete scale and the other one on a continuous scale.

Method 1 (Discrete Rationality Measure)

The first method presented is based on the following point of departure: *a person which has the same respective order of execution state activation levels for the different options compared to the order of world characteristics λ_i will be considered highly rational.* So in this method the rank of the average value λ_i at any given time unit is determined, and compared with the rank of the respective execution state levels. More specifically, the following formula is used to determine the irrationality factor IF .

$$IF = \sum_{i=1}^n \text{abs}(\text{rank}(\text{es}_{a_i}) - \text{rank}(\lambda_i))$$

where n is the number of options available. This irrationality factor tells to which extent the person is behaving rationally in the sense that the higher the irrationality factor IF is, the lower is the rationality of the person. It is assumed that there is uniqueness in ranking and none of the two values assign a similar rank. To calculate the discrete rationality factor DRF , the maximum possible irrationality factor $Max. IF$ can be determined as follows.

$$Max. IF = \frac{n(n+1)}{2} - \text{ceiling}\left(\frac{n}{2}\right)$$

Here $\text{ceiling}(x)$ is the first integer higher than x . Note that Max. IF is approximately $\frac{1}{2}n^2$. As a higher IF means lower rationality, the discrete rationality factor DRF is calculated as:

$$DRF = 1 - \frac{IF}{\text{Max. IF}}$$

On this scale, for each n only a limited number of values are possible; for example, for $n = 3$ three values are possible: 0, 0.5, and 1. In general $\frac{1}{2} \text{Max. IF} + 1$ values are possible, which is approximately $\frac{1}{4}n^2 + 1$. As an example, suppose during a simulation average values of $\lambda_1 = 0.107636$, $\lambda_2 = 0.203044$, and $\lambda_3 = 0.888522$ are given, whereas the execution state values for es_{a_1} , es_{a_2} , and es_{a_3} at a given time point are 0.170554, 0.12367 and 0.43477, respectively. Then according to the given data the world's ranks will be 3, 2, 1 for λ_1 , λ_2 , λ_3 and the person's ranks are 2, 3, 1 for es_{a_1} , es_{a_2} , and es_{a_3} , respectively. So, according to the given formulas $IF = 2$, $\text{Max. IF} = 4$ and $DRF = 0.5$. So in this particular case at this given time point the person is behaving rationally for 50 %.

Method 2 (Continuous Rationality Measure)

The second method presented is based on the following point of departure: a person which receives the maximum benefit will be the highly rational person. This is only possible if es_{a_i} gets 1 for the option whose λ_i is the highest. In this method to calculate the continuous rationality factor CRF , to account for the effort spent in performing actions, the execution state values ES_i for es_{a_i} are normalised as follows.

$$nES_i = \frac{ES_i}{\sum_{i=1}^n ES_i}$$

Here n is the number of options available. Based on this, the continuous rationality factor CRF is determined as follows, with $\text{Max}(\lambda_i)$ the maximal value of the different λ_i .

$$CRF = \frac{\sum_{i=1}^n nES_i \lambda_i}{\text{Max}(\lambda_i)}$$

This method enables to measure to which extent the person is behaving rationally in a continuous manner. For the given example used to illustrate the previous method $CRF = 0.6633$. So according to this method the person is considered to behaving for 66.33 % rationally in the given world. Figures 6.17, 6.18, 6.19 and 6.20 show the two types of rationality (depicted as percentages) of the person for the different scenarios for a changing stochastic world. In these figures the first 250 time points show the rationality achieved by the person just before changing the world characteristics for the simulations shown in Sect. 6.5; after the world change rationality drops drastically, but then starts to increase again in a process to adapt to the changed world.

From time point 250 onwards, the rationality of the person after the change has been made is depicted in Figs. 6.17, 6.18, 6.19 and 6.20 (see Sect. 6.5). The depicted results show that in all four cases the rationality of the person increases over time for the given (new) world. In three of the four cases the discrete rationality notion reaches 100 %. The only exception is the (classical) stimulus-response connection which does not involve the valuing based on the associated feelings. The other connections do involve the feeling association and based on them the person turns out to reach a higher level of rationality according to the discrete rationality measure.

Fig. 6.17 Rationality during learning of the responding connection weight ω_{1i} (A)

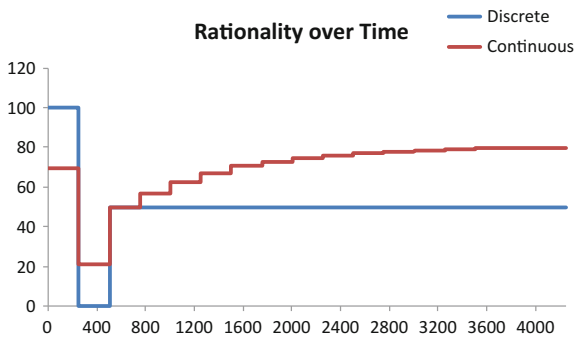


Fig. 6.18 Rationality during learning of the amplifying connection weight ω_{2i} (B)

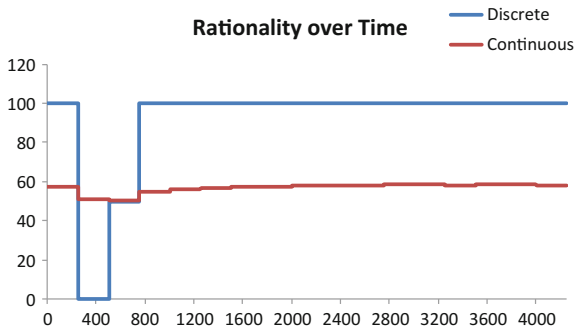


Fig. 6.19 Rationality during learning of the predicting connection weight ω_{3i} (C)

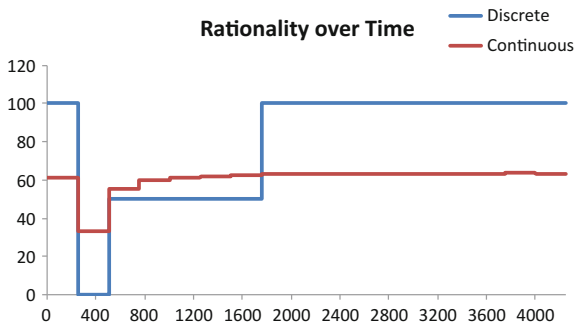
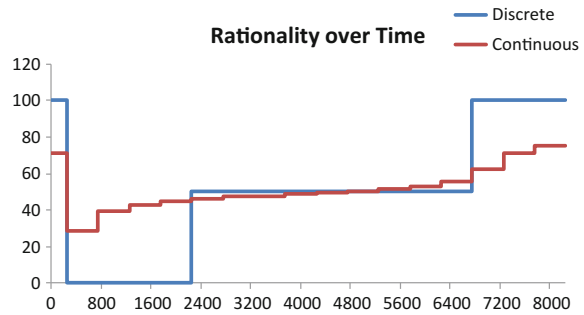


Fig. 6.20 Rationality during learning weights for all three connections ω_{1i} , ω_{2i} , ω_{3i} (ABC)



6.7 Discussion

In this chapter, by a Network-Oriented Modeling approach it was addressed in how far an adaptive human-like temporal-causal network model based on generated emotions for the decision options shows a form of rationality. Parts of this chapter are based on Treur and Umair (2011, 2015). The topic was analysed for a decision model based on valuing of predictions involving feeling states generated in the amygdala; e.g., Bechara et al. (2003), Bechara et al. (1999), Damasio (1994, 2004), Montague and Berns (2002), Janak and Tye (2015), Jenison et al. (2011), Morrison and Salzman (2010), Ousdal et al. (2014), Pessoa (2010), Rangel et al. (2008) and Rudebeck and Murray (2014). To adapt to a specific environment, the model was made adaptive using Hebbian learning; e.g., Gerstner and Kistler (2002) and Hebb (1949).

Two measures were introduced to evaluate the extent of rationality with respect to given world characteristics. Using these measure, the extents of rationality of the different models were analysed. The notions of rationality were chosen in such a manner that the more the person makes the most beneficial choices within the given environment, the more rational it is. It was shown how according to these measures by the adaptivity of the person as modeled by Hebbian learning a high level of rationality was obtained. It was also found that this is robust for major changes in the world: after such a world change, after some delay the rationality level is achieved again. It has been demonstrated that emotion-related valuing of predictions in the amygdala as a basis for adaptive decision making according to Hebbian learning indeed satisfies rationality criteria. Thus the model shows how in human-like decision making emotions serve as a vehicle to obtain rational decisions. This contrasts with the traditional view that emotions and rationality disturb each other in decision processes. Recent findings from neuroscience show that this traditional view is an inadequate way of conceptualisation of processes in the brain, and the model analysed here takes these new insights into account.

The presented model has been extended by social interaction, which shows how the social context also contributes to rationality for the case of collective decision making; see Bosse et al. (2012). In Treur and Umair (2012) it is discussed how some other types of decision and learning models can be evaluated according to the rationality measures used here. In Abro and Treur (2016) an adaptive

temporal-causal network model is described for decision making on food choice in which different valuing perspectives are integrated, among which valuing perspectives related to satisfaction with respect to short term and long term goals.

References

- A.H. Abro, J. Treur, Doubting what to eat: a computational model for food choice using different valuing perspectives (2016)
- A. Bechara, H. Damasio, A.R. Damasio, G.P. Lee, Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *J. Neurosci.* **19**, 5473–5481 (1999)
- A. Bechara, H. Damasio, A.R. Damasio, Role of the amygdala in decision-making. *Ann. NY. Acad. Sci.* **985**, 356–369 (2003)
- T. Bosse, J. Treur, M. Umair, Rationality for adaptive collective decision making based on emotion-related valuing and contagion, in *Modern Advances in Intelligent Systems and Tools, Proceedings of the 25th International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems, IEA/AIE'12, Part II. Studies in Computational Intelligence*, ed. by W. Ding et al., vol. 431 (Springer Verlag, 2012), pp. 103–112
- A. Damasio, *Descartes' Error: Emotion, Reason and the Human Brain* (Papermac, London, 1994)
- A. Damasio, *Looking for Spinoza* (Vintage books, London, 2004)
- A.R. Damasio, *Self Comes to Mind: Constructing the Conscious Brain* (Pantheon Books, NY, 2010)
- W. Gerstner, W.M. Kistler, Mathematical formulations of Hebbian learning. *Biol. Cybern.* **87**, 404–415 (2002)
- H.T. Ghashghaei, C.C. Hilgetag, H. Barbas, Sequence of information processing for emotions based on the anatomic dialogue between prefrontal cortex and amygdala. *Neuroimage* **34**, 905–923 (2007)
- D.O. Hebb, *The Organization of Behaviour* (New York: John Wiley & Sons, 1949)
- P.H. Janak, K.M. Tye, From circuits to behaviour in the amygdala. *Nature* **517**(2015), 284–292 (2015)
- R.L. Jensen, A. Rangel, H. Oya, H. Kawasaki, M.A. Howard, Value encoding in single neurons in the human amygdala during decision making. *J. Neurosci.* **31**(2011), 331–338 (2011)
- J. Kim, *Philosophy of Mind*. (Westview Press, 1996)
- E. Likhtik, R. Paz, Amygdala–prefrontal interactions in (mal)adaptive learning. *Trends Neurosci.* **38**(2015), 158–166 (2015)
- K.A. Lindquist, L.F. Barrett, A functional architecture of the human brain: emerging insights from the science of emotion. *Trends Cogn. Sci.* **16**(2012), 533–540 (2012)
- G. Loewenstein, J. Lerner, The role of emotion in decision making, in *The Handbook of Affective Science*, ed. by R.J. Davidson, H.H. Goldsmith, K.R. Scherer (Oxford University Press, Oxford, England, 2003), pp. 619–642
- P.R. Montague, G.S. Berns, Neural economics and the biological substrates of valuation. *Neuron* **36**, 265–284 (2002)
- S.E. Morrison, C.D. Salzman, Re-valuing the amygdala. *Curr. Opin. Neurobiol.* **20**, 221–230 (2010)
- E.A. Murray, The amygdala, reward and emotion. *Trends Cogn. Sci.* **11**, 489–497 (2007)
- O.T. Ousdal, K. Specht, A. Server, O.A. Andreassen, R.J. Dolan, J. Jensen, The human amygdala encodes value and space during decision making. *Neuroimage* **101**(2014), 712–719 (2014)
- J.M. Pearson, K.K. Watson, M.L. Platt, Decision making: the neuroethological turn. *Neuron* **82** (2014), 950–965 (2014)
- L. Pessoa, Emotion and cognition and the amygdala: from “what is it?” to “what’s to be done?”. *Neuropsychologia* **48**(2010), 3416–3429 (2010)

- A. Rangel, C. Camerer, P.R. Montague, A framework for studying the neurobiology of value-based decision making. *Nat. Rev. Neurosci.* **9**, 545–556 (2008)
- P.H. Rudebeck, E.A. Murray, The orbitofrontal oracle: cortical mechanisms for the prediction and evaluation of specific behavioral outcomes. *Neuron* **84**(2014), 1143–1146 (2014)
- C.D. Salzman, S. Fusi, Emotion, cognition, and mental state representation in amygdala and prefrontal cortex. *Annu. Rev. Neurosci.* **33**, 173–202 (2010)
- L.P. Sugrue, G.S. Corrado, W.T. Newsome, Choosing the greater of two goods: neural currencies for valuation and decision making. *Nat. Rev. Neurosci.* **6**, 363–375 (2005)
- J. Treur, M. Umair, Emotions as a Vehicle for Rationality: Rational Decision Making Models Based on Emotion-Related Valuing and Hebbian Learning. *Biologically Inspired Cognitive Architectures* **14**, 40–56 (2015)
- J. Treur, M. Umair, On rationality of decision models incorporating emotion-related valuing and Hebbian learning, in *Proceedings of the 18th International Conference on Neural Information Processing, ICONIP'11, Part III. Lecture Notes in Artificial Intelligence*, eds. by B.-L. Lu, L. Zhang, J. Kwok, vol. 7064 (Springer Verlag, 2011), pp. 217–229
- J. Treur, M. Umair, Rationality for temporal discounting, memory traces and Hebbian learning, in *Modern Advances in Intelligent Systems and Tools, Proceedings of the 25th International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems, IEA/AIE'12, Part II. Studies in Computational Intelligence*, eds. by W. Ding et al., vol. 431 (Springer Verlag, 2012), pp. 53–61

Part III
Youself and the Others

Chapter 7

From Mirroring to the Emergence of Shared Understanding and Collective Power

Biological and Computational Perspectives on the Emergence of Social Phenomena

Abstract Shared understanding and collective power are social phenomena that serve as a form of glue between individual persons. They easily emerge and often involve both cognitive and affective aspects. As the behaviour of each person is based on complex internal mental processes involving, for example, own goals, emotions and beliefs, it would be expected that such sharedness and collectiveness is very hard to achieve. Apparently, specific mechanisms are required to tune the individual mental processes to each other in order to enable the emergence of shared mental states and collective behaviour. Having knowledge about these mechanisms from a biological context provides a basis to modeling corresponding mechanisms in a computational setting. From a biological perspective, mirror neurons and internal simulation are core concepts to explain the mechanisms underlying such social phenomena. In this chapter it is discussed how through a Network-Oriented Modeling approach such neurological concepts can be used to obtain human-like temporal-causal network models for such social phenomena. It is discussed how these models indeed are an adequate basis for the emergence of social phenomena such as shared understanding and collective power.

7.1 Introduction

In society often some form of ‘sharedness’ of understanding or ‘collectiveness’ of action is experienced, which usually covers both cognitive and affective dimensions. Although this is a very common type of social phenomenon, at forehand it is not at all clear how it can emerge. For example, the experience of feeling good being part of a group with a shared understanding and collective action may be experienced as quite natural. However, as persons in a group are autonomous agents with their own neurological structures and patterns, carrying, for example, their own emotions, beliefs, desires and intentions, it would be more reasonable to expect that such sharedness and collectiveness is impossible to achieve. Nevertheless, often groups

develop coherent views and decisions, and, even more surprisingly, the group members seem to share a positive feeling with it. These processes depend on possibilities for informational and emotional transfer between individuals, which can be enhanced by technological infrastructure such as social media.

In recent years by developments in neuroscience new light has been shed on this seeming paradox of individuality versus sharedness and collectiveness. This has led to the new discipline called Social Neuroscience; e.g., Cacioppo and Berntson (2005), Cacioppo et al. (2006), Decety and Cacioppo (2010), Decety and Ickes (2009), Harmon-Jones and Winkielman (2007). Two interrelated core concepts in this discipline are mirror neurons and internal simulation of another person's mental processes. Mirror neurons are neurons that not only have the function to prepare for a certain action or body change, but are also activated upon observing somebody else who is performing this action or body change; e.g., Iacoboni (2008), Pineda (2009), Rizzolatti and Sinigaglia (2008). Internal simulation is internal mental processing that copies processes that may take place externally, for example, mental processes in another individual; e.g., Damasio (1994, 1999), Gallese and Goldman (1998), Goldman (2006), Hesslow (1994, 2002, 2012). Mechanisms involving these core concepts have been described that provide an explanation of the emergence of sharedness and collectiveness from a biological perspective. Formalisation of such mechanisms provides a basis for the design of human-like temporal-causal network models, able to show similar patterns of emerging shared understanding and collective action.

The type of biologically inspired human-like temporal-causal network models discussed in this chapter have a number of possible application areas. In the first place they can be used for the analysis of human social processes in groups, crowds or in societies as a whole. Examples of this are the analysis of collective decision making in groups, crowd behaviour in emergency situations, social contagion of emotions and opinions, and the development of societal or political movements. A second area of application addresses analysis and design of socio-technological systems including, for example, social media. This concerns groups that partly consist of humans and partly of artificial agents in the form of devices such as smartphones, and the use of human-made infrastructure for communication such as Internet and social media. For such socio-technological systems it can not only be analysed what patterns may emerge under which circumstances, but the focus can also be on the design of these devices and media, in order to create a situation in which the right types of patterns emerge, for example, to enable safe evacuation in emergency situations. A third possible area of application concerns close empathic interaction and emotional attachment between a human and a device.

In this chapter, first in Sect. 7.2 the concepts of mirror neurons and internal simulation as mentioned are briefly reviewed, and it is discussed how they can be modeled by a Network-Oriented Modeling approach. Next, in Sect. 7.3 it is discussed how based on biological mechanisms involving these concepts shared understanding can emerge. This covers both cognitive and affective understanding, and in a combined form empathic understanding. In Sect. 7.4 biological mechanisms are discussed enabling the emerge of collective decisions and actions, and it is shown

how such collective actions can be grounded in shared cognitive and affective understanding. Section 7.5 illustrates how in general the integration of internal and external impacts can be modeled as computational mechanisms in temporal-causal network models, and in Sect. 7.6 it is pointed out how certain abstraction methods can be applied to such models. Finally, Sect. 7.7 is a discussion.

7.2 Mirror Neuron Activation and Internal Simulation

Social contagion is the basic phenomenon that due to social interaction persons affect each other's states and behaviours, such as emotions, beliefs, opinions, preferences or appreciations, eating and drinking behaviour, and so on. Sometimes sayings refer to this, for example the saying 'Mom happy, everybody happy', which assumes a central role for the mom in the social interaction patterns in the family, and as a consequence due to contagion also in the family members' happiness. Social contagion could be addressed abstracting from internal processes within the persons. An assumption can be made that contagion takes place, but the question about which internal mechanisms within persons make this contagion happen may still leave an unsatisfied feeling. In how far can a model not taking into account these internal mechanisms be human-like? In the current section these mechanisms will be discussed, as they have been discovered in recent years: mirror neurons and internal simulation of another person's mental processes. Together these concepts are a basis for biological mechanisms that realise an individual's mental function of mirroring mental processes of another individual, and that mechanism explains the assumption on social contagion. More specifically, this function plays a crucial role in enabling the emergence of shared and empathic understanding and collective action, as will be discussed in Sects. 7.3 and 7.4.

7.2.1 *The Discovery of Mirror Neurons*

Recently it has been found that in humans a specific type of neurons exists, called *mirror neurons*, which both are active when the person prepares for certain actions or bodily changes and when the person observes such actions or body states of other persons. The discovery of mirror neurons originates from single cell recording experiments with monkeys in Parma in the 1990s. In particular, the focus was on an area in the premotor cortex (F5) involved in the preparation of grasp actions. To their own surprise, the researchers discovered that some of the recorded cells were not only firing when the monkey was preparing a grasp action, but also when somebody in the lab was grasping something and the monkey just observed that, see Gallese et al. (1996), Rizzolatti et al. (1996); see also Iacoboni (2008a), Rizzolatti and Craighero (2004), Rizzolatti and Sinigaglia (2008). The highly unexpected element was that sensory processing of observed actions of others

involves neurons that are also involved in the subject's preparation for the same type of action. Traditionally sensory processing was assumed to be separate from preparing. It turned out that in the premotor area F5 about 20 % of the neurons are both active when preparing and when observing the action.

After the discovery of mirror neurons in monkeys it has been hypothesized that similar types of neurons also occur in humans. Indeed, for humans from the usual imaging methods it can be found that in certain premotor areas activity occurs both when an action is observed and when the action is prepared; e.g., Cochin et al. (1999), Gastout and Bert (1954) based on EEG data, (Grafton et al. 1996; Rizzolatti et al. 1996) based on PET data, and (Iacoboni et al. 2005) based on fMRI; see also reviews in Molenberghs et al. (2012), Kilner and Lemon (2013). However, due to limitations in resolution, from such methods it cannot be found whether the neurons active in action observation are exactly the same neurons as those that are active in preparing for an action. In principle they could be different neurons in the same area. Therefore in the years after the discovery of mirror neurons in monkeys it still has been subject to debate whether they also exist in humans; e.g., Hickok (2009). But in recent years the existence of mirror neurons in humans has found support in single cell experiments with epilepsy patients undergoing pre-surgical evaluation of the foci of epilepsy; e.g., Fried et al. (2011), Mukamel et al. (2010); see also Iacoboni (2008a, pp. 201–203) and Iacoboni (2008b), Keysers and Gazzola (2010). In these experiments for 14 patients the activity of approximately 500 neurons was recorded; they were located in three sectors of the mesial frontal cortex (the ventral and dorsal sectors of the anterior cingulate cortex and the pre-supplementary motor cortex (SMA)/SMA proper complex). The subjects were tested both for hand-grasping actions and for emotional face expressions. Some of the main findings were that neurons with mirror neuron properties were found in all sites in the mesial frontal cortex where recording took place, in total for approximately 12 % of all recorded neurons; about half of them related to hand-grasping, and the other half to emotional face expressions; e.g., Iacoboni (2008b).

7.2.2 Neurons for Control and Self-other Distinction

Due to the multiple functions of mirror neurons, the functional meaning of activation of them (e.g., preparing or observing an action, or both) in principle is context-dependent. The context determines in which cases their activation is meant to lead to actual execution of the action (e.g., in self-initiated action performance, or imitation), and in which cases it is not (e.g., in action observation and interpretation). A specific set of neurons has been found (sometimes called super mirror neurons: Iacoboni (2008b)) that seem to be able to indicate such a context and play a role in the control of actual execution of a prepared action. These neurons are suggested to exert control by allowing or suppressing action execution and/or preparation states. More specifically, in the single cell recording experiments with

epileptic patients mentioned above, also cells were found that are active when the person prepares an own action that is executed, but shut down when the action is only observed. This has led to the hypothesis that these cells may be involved in the functional distinction between a preparation state activated in order to actually perform the action, or a preparation state only activated to interpret an observed action. In Iacoboni (2008a, pp. 201–202) it is also described that some of such cells are sensitive to a specific person, so that the action can be attributed to the specific person that was observed: self-other distinction; see also Brass and Spengler (2009). More details on such types of neurons can be found in Brass and Spengler (2009), Iacoboni (2008a, b, pp. 196–203).

7.2.3 *Generating Emotions and Feelings by Internal Simulation: As-if Body Loops*

Activation of mirror neurons is important not by itself, but because it plays a crucial role in an important mental function: *mirroring* mental processes of other persons by *internal simulation*. How mirroring relates to internal processes involving emotions and feelings may ask for some further explanation.

Recall from Chap. 3, Sect. 3.2 the discussion about the classical view on emotions assuming that based on some sensory input, due to internal processing associated emotions are felt, and subsequently these emotions are expressed in some body state; e.g., a face expression. As an alternative for this classical view the notions of body loop and of as-if body loop were discussed. An as-if body loop is a form of internal simulation of the own body states; this is a process in which the preparation for some body state affects the sensory representation of this body state, which in turn affects the preparation of this body state (see Chap. 3, Sect. 3.2):

sensory representation of $s \rightarrow$ preparation for body state $b \rightarrow$

sensory representation of (simulated) body state $b \rightarrow$ preparation for body state b

This provides a cyclic process that (for a constant environment) can lead to equilibrium states for feelings (based on sensory representations of body states) and emotional responses.

7.2.4 *Mirroring Process: Mirror Neuron Activation and Internal Simulation*

From a more general viewpoint, as-if body loops as introduced by Damasio (1994) contribute:

- (1) sensory input directly affects preparation states and execution states, after which further internal processing takes place, in line with, e.g., James (1884)

- (2) the notion of internal simulation involving preparations and sensory representations, in line with, e.g., Becker and Fuchs (1985, Hesslow (1994, 2002, 2012), Barsalou (2009), Marques and Holland (2009), Pezzulo et al. (2013).

Here (1) breaks with the tradition that there is a standard order of processing sensing—internal processing—preparation for action, and (2) allows for involving changing body representations in internal processes without actually having to change any body state. As mirror neurons make that some specific sensory input (an observed person) directly links to related preparation states, just like (1) above, it fits quite well in the perspective based on as-if body loops. In this way mirroring is a process that fully integrates mirror neuron activation states in the ongoing internal simulation processes based on as-if loops; see also Damasio (2010, pp. 102–104), or Chap. 3.

From a behavioural perspective social contagion processes can be described at a conceptual level by a graphical conceptual representation of a temporal-causal network model as shown in Fig. 7.1 (in this case for two persons only); see Table 7.1 for notations and explanations for the states used. For a person Y behaviour b (denoted by $es_{Y,b}$) can be affected by some stimulus s , but also by sensing the same behaviour shown by another person. In the latter case the (executed) behaviour b (for example, showing a face expression with a specific emotion) of one person X is sensed by another person Y , through Y 's sensor state $ss_{Y,es_{X,b}}$ and vice versa. Note that now different persons are involved, always an extra indication (subscript) is needed in the notations to make clear which person's state is meant.

The graphical conceptual representation shown in Fig. 7.1 can be described by a conceptual matrix representation as shown in Table 7.2.

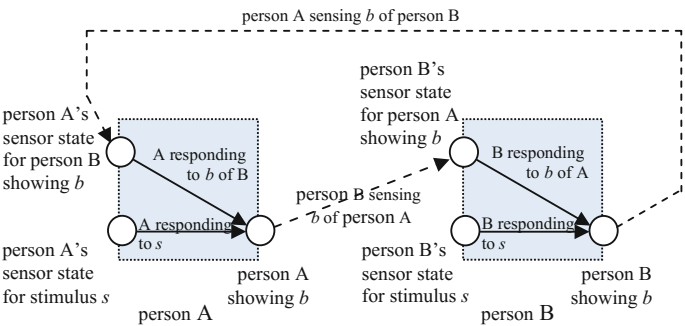


Fig. 7.1 Conceptual representation of a behavioural temporal-causal network model for social contagion

Table 7.1 Notations and explanations for the states used in the conceptual representation of the behavioural model depicted in Fig. 7.1

State	Description
$ss_{Y,s}$	Sensor state of person Y for stimulus s
$ss_{Y,es_{X,b}}$	Sensor state of person Y sensing execution state $es_{X,b}$ of person X for b
$es_{Y,b}$	Execution state of person Y for b

Table 7.2 Conceptual matrix representation for the behavioural model depicted in Fig. 7.1

To from	$SS_{Y,s}$	$es_{Y,b}$	$SS_{Y,es_{X,b}}$
$SS_{Y,s}$		$\omega_{Y,responding_to_s}$	
$es_{X,b}$			$\omega_{Y,sensing_b_of_X}$
$SS_{Y,es_{X,b}}$		$\omega_{Y,responding_to_b_of_X}$	
η_Y	—	$\eta_{es_{Y,b}}$	$\eta_{SS_{Y,es_{X,b}}}$
$c_Y(\dots)$	—	$c_{es_{Y,b}}(\cdot)$	$c_{SS_{Y,es_{X,b}}}(\cdot)$

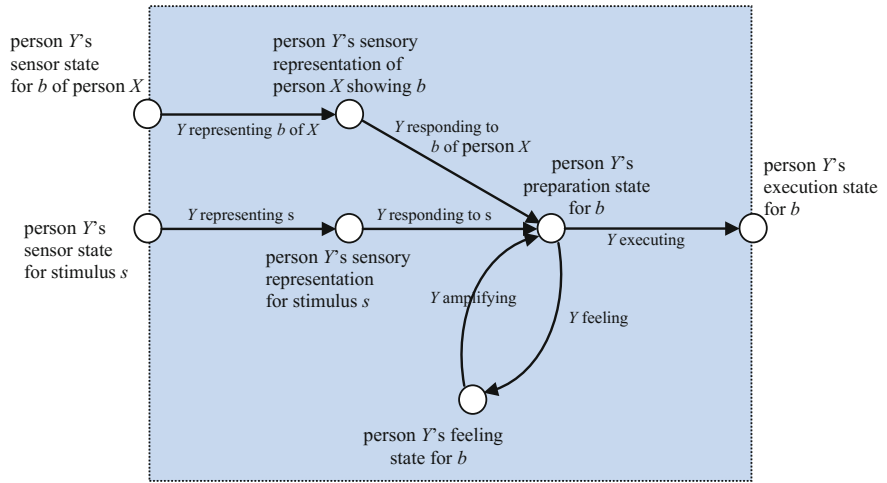


Fig. 7.2 Graphical conceptual representation of a temporal-causal network model for mirroring process within a person Y responding to another person X: mirror neuron activation and internal simulation by an as-if body loop

This behavioural model only provides a kind of black box view on the contagion processes. The mirroring process is the neural mechanism within the persons realising the contagion processes; this is shown in a more detailed view on the internal processes as depicted in Fig. 7.2; see Table 7.2 for notations and explanations for the states used. This graphical conceptual representation can be seen as a refinement (zooming in) of what is depicted as a box in the graphical representation in Fig. 7.1 (Table 7.3).

Here, again the preparation for body state *b* (e.g., some emotional response) can either be triggered by sensing an external stimulus *s* (path in the middle of the graph shown in Fig. 7.2), or by observing somebody else performing *b* (upper path in the graph shown in Fig. 7.2). However, now in both cases, as a first step, from the sensor states sensory representations are generated and these sensory representations in turn affect the preparation state. Moreover, after generation of the preparation state, internal simulation takes place based on the as-if body loop (lower part

Table 7.3 Notations and explanations of the states in the internal temporal-causal network model depicted in Fig. 7.2

State	Description
$ss_{Y,s}$	Sensor state of person Y for stimulus s
$es_{Y,b}$	Execution state of person Y for b
$ss_{Y,es_{X,b}}$	Sensor state of person Y for execution state of person X for b
$srs_{Y,s}$	Sensory representation of person Y for stimulus s
$srs_{Y,es_{X,b}}$	Sensory representation of person Y for execution state of person X for b
$ps_{Y,b}$	Preparation state of person Y for b
$fs_{Y,b}$	Feeling state of person Y for b

Table 7.4 Conceptual matrix representation for the internal temporal-causal network model depicted as a graphical representation in Fig. 7.2

To from	$ss_{Y,s}$	$es_{Y,b}$	$ss_{Y,es_{X,b}}$	$srs_{Y,s}$	$srs_{Y,es_{X,b}}$	$ps_{Y,b}$	$fs_{Y,b}$
$ss_{Y,s}$				$\omega_{Y,representing_s}$			
$es_{X,b}$			$\omega_{Y,sensing_b_of_X}$				
$ss_{Y,es_{X,b}}$					$\omega_{Y,representing_b_of_X}$		
$srs_{Y,s}$						$\omega_{Y,responding_to_s}$	
$srs_{Y,es_{X,b}}$						$\omega_{Y,responding_to_b_of_X}$	
$ps_{Y,b}$		$\omega_{Y,executing}$					$\omega_{Y,feeling}$
$fs_{Y,b}$						$\omega_{Y,amplifying}$	
η_Y	–	$\eta_{es_{Y,b}}$	$\eta_{ss_{Y,es_{X,b}}}$	$\eta_{srs_{Y,s}}$	$\eta_{srs_{Y,es_{X,b}}}$	$\eta_{ps_{Y,b}}$	$\eta_{fs_{Y,b}}$
$c_Y(\dots)$	–	$c_{es_{Y,b}}(\cdot)$	$c_{ss_{Y,es_{X,b}}}(\cdot)$	$c_{srs_{Y,s}}(\cdot)$	$c_{srs_{Y,es_{X,b}}}(\cdot)$	$c_{ps_{Y,b}}(\cdot)$	$c_{fs_{Y,b}}(\cdot)$

in Fig. 7.2) which affects both the activation levels of the related feeling and the preparation state. Finally based on the activation level of the preparation state the corresponding execution state can be activated by which the behaviour is shown. Note that, as this mirroring process happens mostly in an unconscious manner, in a social context mirroring imposes (often unnoticed) serious limitations on the freedom for individuals to have their own personal emotions, beliefs, intentions, and actions. Table 7.4 shows a conceptual matrix representation of the model depicted as a graphical conceptual representation in Fig. 7.2.

A numerical representation for the model depicted by the conceptual representation in Table 7.4 is shown in Box 7.1

For persons X and Y with $X = A$ and $Y = B$, or $X = B$ and $Y = A$:

LP1 Person Y sensing b of person X

$$\begin{aligned} dss_{Y,es_{X,b}}/dt &= \eta_{ss_{Y,es_{X,b}}} [c_{ss_{Y,es_{X,b}}}(\omega_{Y,sensing_b_of_X}es_{X,b}) - ss_{Y,es_{X,b}}] \\ ss_{Y,es_{X,b}}(t + \Delta t) &= ss_{Y,es_{X,b}}(t) + \eta_{ss_{Y,es_{X,b}}} [c_{ss_{Y,es_{X,b}}}(\omega_{Y,sensing_b_of_X}es_{X,b}(t)) - ss_{Y,es_{X,b}}(t)]\Delta t \end{aligned}$$

LP2 Person Y representing b of person X

$$\begin{aligned} \text{dsrs}_{Y,es_{X,b}}/\text{dt} &= \eta_{\text{srs}_{Y,es_{X,b}}} [c_{\text{srs}_{Y,es_{X,b}}}(\omega_{Y,\text{representing_}b_of_X\text{SS}_{Y,es_{X,b}}}) - \text{srs}_{Y,es_{X,b}}] \\ \text{srs}_{Y,es_{X,b}}(t + \Delta t) &= \text{srs}_{Y,es_{X,b}}(t) \\ &\quad + \eta_{\text{srs}_{Y,es_{X,b}}} [c_{\text{srs}_{Y,es_{X,b}}}(\omega_{Y,\text{representing_}b_of_X\text{SS}_{Y,es_{X,b}}}(t)) - \text{srs}_{Y,es_{X,b}}(t)] \Delta t \end{aligned}$$

LP3 Person Y representing stimulus s

$$\begin{aligned} \text{dsrs}_{Y,s}/\text{dt} &= \eta_{\text{srs}_{Y,s}} [c_{\text{srs}_{Y,s}}(\omega_{Y,\text{representing_}s\text{SS}_{Y,s}}) - \text{srs}_{Y,s}] \\ \text{srs}_{Y,s}(t + \Delta t) &= \text{srs}_{Y,s}(t) + \eta_{\text{srs}_{Y,s}} [c_{\text{srs}_{Y,s}}(\omega_{Y,\text{representing_}s\text{SS}_{Y,s}}(t)) - \text{srs}_{Y,s}(t)] \Delta t \end{aligned}$$

LP4 Person Y generating and amplifying response b

$$\begin{aligned} \text{dps}_{Y,b}/\text{dt} &= \eta_{\text{ps}_{Y,b}} [c_{\text{ps}_{Y,b}}(\omega_{Y,\text{responding_to_}s\text{Srs}_{Y,s}}, \omega_{Y,\text{responding_to_}b_of_X\text{Srs}_{Y,es_{X,b}}}, \\ &\quad \omega_{Y,\text{amplifying fs}_{Y,b}}) - \text{ps}_{Y,b}] \\ \text{ps}_{Y,b}(t + \Delta t) &= \text{ps}_{Y,b}(t) + \eta_{\text{ps}_{Y,b}} [c_{\text{ps}_{Y,b}}(\omega_{Y,\text{responding_to_}s\text{Srs}_{Y,s}}(t), \\ &\quad \omega_{Y,\text{responding_to_}b_of_X\text{Srs}_{Y,es_{X,b}}}(t), \omega_{Y,\text{amplifying fs}_{Y,b}}(t)) - \text{ps}_{Y,b}(t)] \Delta t \end{aligned}$$

LP5 Person Y feeling b

$$\begin{aligned} \text{dfs}_{Y,b}/\text{dt} &= \eta_{\text{fs}_{Y,b}} [c_{\text{fs}_{Y,b}}(\omega_{Y,\text{feeling ps}_{Y,b}}) - \text{fs}_{Y,b}] \\ \text{fs}_{Y,b}(t + \Delta t) &= \text{fs}_{Y,b}(t) + \eta_{\text{fs}_{Y,b}} [c_{\text{fs}_{Y,b}}(\omega_{Y,\text{feeling ps}_{Y,b}}(t)) - \text{fs}_{Y,b}(t)] \Delta t \end{aligned}$$

LP6 Person Y executing b

$$\begin{aligned} \text{des}_{Y,b}/\text{dt} &= \eta_{\text{es}_{Y,b}} [c_{\text{es}_{Y,b}}(\omega_{Y,\text{executing ps}_{Y,b}}) - \text{es}_{Y,b}] \\ \text{es}_{Y,b}(t + \Delta t) &= \text{es}_{Y,b}(t) + \eta_{\text{es}_{Y,b}} [c_{\text{es}_{Y,b}}(\omega_{Y,\text{executing ps}_{Y,b}}(t)) - \text{es}_{Y,b}(t)] \Delta t \end{aligned}$$

The symbols are explained as follows:

s	external stimulus
b	behavioural response, for example, an emotion
$\text{es}_{Y,b}$	execution state: person Y shows response b
$\text{ss}_{Y,Z}$	sensor state: person Y senses Z with $Z = s$ or $Z = \text{es}_{X,b}$
$\text{ps}_{Y,b}$	preparation state: person Y prepares for response b
$\text{es}_{Y,b}$	execution state: person Y shows response b
$\text{fs}_{Y,b}$	feeling state of person Y for b
$\text{srs}_{Y,Z}$	sensory representation of person Y for Z with $Z = s$ or $Z = \text{es}_{X,b}$
$\text{ss}_{Y,Z}$	person Y senses Z with $Z = s$ or $Z = \text{es}_{X,b}$
η_S	update speed parameter for state S
c_S	combination function for state S

Box 7.1 Numerical representation of the temporal-causal network model shown as conceptual matrix representation in Table 7.4 social contagion integrating internal mirroring processes

Note that the connection weights of some of the connections represent aspects of the relation between person X and Y . For example, the weights of three connections determine *how open or responsive* person Y is for influence from person X :

- the weight $\omega_{Y,\text{sensing_}b_of_X}$ of the sensing connection from the execution state of person X to the sensor state of person Y
- the weight $\omega_{Y,\text{representing_}b_of_X}$ of Y 's representation connection from Y 's sensor state to Y 's sensory representation state
- the weight $\omega_{Y,\text{responding_}b_of_X}$ of Y 's responding connection from Y 's sensory representation state to Y 's preparation state

If one or more of these connections have very low weights, then person Y is hardly affected by person X . Moreover, the overall effect of X on Y also depends on the strength $\omega_{X,\text{executing}}$ of the connection within person X from preparation state to execution state (person X 's *expressiveness*). If person X is not expressing very well, then person Y does not sense much to respond to. The specific weights of such connections represent specific characteristics of persons and relationships. By assuming variations in weights for these connections, a wide variety of persons and relationships can be modeled.

More complex variations of social contagion models involving both beliefs and emotions or intentions and emotions have been addressed in Hoogendoorn et al. (2011), using similar mechanisms as discussed here. Moreover, in Bosse et al. (2013) the model ASCRIBE is presented, in which internal dynamics involving emotions, beliefs and intentions and the social contagion of these states are integrated in a similar manner; also see Sect. 7.5.

7.2.5 *Development of the Discipline Social Neuroscience*

Above it has been pointed out how mirroring internal processes and states of other persons lead to activation of a person's corresponding own processes and states. The latter processes and states at the same time play a crucial role in the person's own feelings and actions. Metaphorically spoken, mirroring has not been designed as a separate mental function with data structures and processes fully disjoint from the other mental functions; instead mirroring is fully integrated in the person's own mental processes and uses shared data structures and processes. This integration provides an effective mechanism for how actions and feelings of other persons and own actions and feelings affect each other. This biological mechanism explains how in a social context persons fundamentally tune their personal actions and states to each other, including their feelings. Given these implications, the discovery of mirror neurons and how they play their role in mirroring processes is considered a crucial step for the further development of the disciplines of social cognition and social psychology, by providing a biological basis for social phenomena. Many examples of social phenomena now can be explained by relating them to biological mechanisms that realize mirroring, for example:

- social diffusion or contagion of personal states (e.g., opinions or emotions)
- empathic understanding
- group formation, group cohesion
- collective decision making.

Based on these developments, and their wide applicability the new discipline Social Neuroscience has shown a fast development; e.g., Cacioppo and Berntson (2005), Cacioppo et al. (2006), Decety and Cacioppo (2010), Decety and Ickes (2009), Harmon-Jones and Winkielman (2007). The impact of this discipline is very wide, as it is considered to cover not only the items indicated above, but also, for example, the concept of social reality (Butz 2008), spiritual and religious experience (Seybold 2010), and collective consciousness or global empathy and its role in the future evolution (Combs and Krippner 2008; Rifkin 2010). In the next two sections it will be discussed in some more detail how different types of shared understanding and collective power can emerge based on mirroring processes.

7.3 The Emergence of Shared Understanding

Understanding can be viewed as a relation between an internal mental state and the world state to which the understanding refers. It can occur in different types, that can be distinguished from each side of the relation: from the internal mental state side, and from the side of the concerning world state to which the understanding refers. First distinctions from the former (internal state) side are discussed, and next distinctions from the latter (world state) side.

A person can have an understanding of a world state by generating and maintaining an internal cognitive state in relation to it (e.g., one or more beliefs about it). This can be distinguished as a *cognitive type of understanding*. A person can also form and maintain an internal affective state in relation to a world state (e.g., a specific emotion or feeling associated to it). Such a form of understanding can be distinguished as an *affective type of understanding*. An important role of this type of understanding is that it provides a basis for *experiencing* in the understanding. Affective and cognitive understanding are often related to each other. Within a person, any cognitive state relates to an associated emotional response which based on an as-if body loop involving a sensory representation of a body state which is the basis of the related feeling (e.g., Damasio 1994, 1999, 2003, 2010); see also Sect. 7.2. When mirroring takes place for both the cognitive and affective state, this provides a mechanism to obtain shared understanding integrating cognitive and affective aspects. For the case of similar neural architectures, the (bidirectional) associations between cognitive state and emotion in an observing person are similar to these associations in an observed person. This will further strengthen the shared integrated cognitive and affective understanding; more extreme cases of this occur in identical twins.

A second way of distinguishing different types of understanding is by considering differentiations of the concerning world state to which the understanding

refers. The world can be conceptualised as a kind of landscape in which persons occur as active, living entities. The internal processes of the persons are also part of the world, and can be distinguished from person-independent aspects of the landscape. Given this picture, understanding may refer to either a *person-external* world state or a *person-internal* world state. For example, having beliefs about another person’s emotions, beliefs or goals is of the second, person-internal type, whereas having beliefs about the weather is of the first type.

The two dimensions of distinctions for types of understanding introduced above can be applied to *shared* understanding of an person B with an person A, from which a matrix results as illustrated in Table 7.5 with different examples.

7.3.1 The Emergence of Shared Understanding for External World States

A person’s understanding of the external world in the form of a collection of beliefs is sometimes called the person’s world model. This can be considered a cognitive world model. More general, shared understanding of an external world state can involve (see Table 7.5):

- a *shared cognitive world model*
(e.g., sharing beliefs about an external world state)
- a *shared affective world model*
(e.g., sharing feelings about an external world state)
- a combined *shared cognitive-affective world model*
(e.g., sharing both beliefs and feelings about an external world state)

Some examples of the first, cognitive type of shared understanding of person-external states for different (sub) populations concerning climate change and its cause are shown in Table 7.6. For example, an understanding shared in population C is that the climate is changing, but this is not due to human action, it has a natural cause. As another example, a shared understanding in population D is that the climate

Table 7.5 Examples of different types of shared understanding

	Person-internal	Person-external
Shared cognitive understanding	• Having beliefs about person A’s beliefs, intentions or goals • Sharing goals for an internal person state	• Sharing beliefs with person A about an external world state • Sharing goals for an external world state
Shared affective understanding	• Feeling the same as person A is feeling about an internal state	• Sharing a good or bad feeling about an external world state
Shared cognitive and affective understanding	• Believing that person A feels bad • Believing X and feeling Y, and believing that person A also believes X and feels Y	• Sharing a belief or goal and feeling • Sharing a belief and a feeling that intention X will achieve goal Y

Table 7.6 Cognitive understanding of person-external states shared in different populations

	Belief on climate	Belief on cause
Population A	Climate is changing	Due to human action
Population B	Climate is not changing	Not due to human action
Population C	Climate is changing	Not due to human action
Population D	Climate is not changing	Due to human action

is not changing, and this is due to human action (e.g., without human action the temperature would have a decreasing trend, but human action compensates for this).

An example of a shared combined cognitive-affective person-external understanding is sharing a belief that climate change has some serious effects and sharing a bad feeling about that, or sharing a belief that a new iPhone will come out soon and sharing a good feeling about that. Obtaining such shared understanding of the external world may make use of different means. Individual information gathering can play a role, but also verbal and nonverbal interaction between persons. If some external world state is considered by persons, both verbal and nonverbal expressions are input for mirroring processes. These mirroring processes affect, for example, both the strength by which something is believed about this state, and the strength of the feeling associated to it. Thus both cognitive and affective shared understanding can develop, based on (mostly unconscious) mirroring processes.

7.3.2 *The Emergence of Shared Understanding for Internal Mental States*

A second type of understanding concerns world states that are internal for one of the persons in the world. For such understanding different terms are used; e.g., mindreading, Theory of Mind (ToM), empathy, or more specific terms such as emotion or intention recognition; e.g., Decety and Ickes (2009), Goldman (2006), Preston and de Waal (2002). Also here understanding may be limited to cognitive understanding; for example, believing that another person has the intention to go out for a dinner, or believing that this person feels depressed. However, for humans also an affective type of mutual understanding is common, usually combined with some form of cognitive understanding. One of the most fundamental forms of mutual understanding is indicated by the notion of *empathy*; e.g., see De Vignemont and Singer (2006), Decety and Ickes (2009), Iacoboni (2008a), Preston and de Waal (2002), Shamay-Tsoory (2008, 2011), Singer and Leiberg (2009). Originally by Lipps (1903) the notion was named by the German word ‘*einfühlung*’ which could be translated as ‘feeling into’; e.g., Preston and de Waal (2002). As this word indicates more explicitly, the notion of empathy has a strong relation to feeling: *empathic understanding* includes experiencing what the other person feels, but also believing that the experienced feeling is felt by the other person, based on self-other distinction. Therefore empathic understanding can be considered a form of

combined affective and cognitive understanding; see also Shamay-Tsoory (2008, 2011). As an example, in Singer and Leiberg (2009), and De Vignemont and Singer (2006, p. 435), the following four criteria of empathy of *B* for *A* are formulated:

- (1) Presence of an affective state in a person *B*
- (2) Isomorphism of *B*'s own and *A*'s affective state
- (3) Elicitation of the *B*'s affective state upon observation of *A*'s affective state
- (4) Knowledge of *B* that *A*'s affective state is the source of the *B*'s own affective state.

The understanding indeed is both affective (1) and cognitive (4), but in this case it concerns in particular an affective state and not a cognitive state of the other person. Therefore it can be called *affective-focused empathy*. In contrast, to indicate affective and cognitive understanding of another person's cognitive state (e.g., another person's belief) the term *cognitive-focused empathy* may be used. The term *full empathy* can be used to indicate combined cognitive-affective understanding of both cognitive and (associated) affective states of another person. Note that empathy always involves feelings, so this is also the case, for example, in cognitive-focused empathy. However, in case of full empathy these feelings are related to the other person (using self-other distinction), and in case of purely cognitive-focused empathy the feelings are experienced, but not related to the other person (for example, due to impaired self-other distinction). Table 7.7 illustrates these types of understanding for person *B* having understanding of states of person *A*. That mirroring (together with self-other distinction) provides a basic mechanism involved in the creation of empathic understanding has much support in the recent literature; e.g., Gallese (2003), Shamay-Tsoory (2008, 2011), Singer and Leiberg (2009), Iacoboni (2008a, pp. 106–129).

Table 7.7 Examples of different types of theory of mind and empathy of person *B* w.r.t. person *A*

Person A Person B	Affective states	Cognitive states	Affective and cognitive states
Affective understanding	Feeling but not having a belief for <i>A</i> 's emotion (<i>emotion contagion</i>)	Feeling but not having a belief for <i>A</i> 's belief	Feeling but not having a belief for <i>A</i> 's emotion and belief
Cognitive understanding	Having a belief but no feeling for <i>A</i> 's emotion (<i>affective-focused ToM</i>)	Having a belief but no feeling for <i>A</i> 's belief (<i>cognitive-focused ToM</i>)	Having a belief but no feeling for <i>A</i> 's emotion and belief (<i>ToM</i>)
Affective and cognitive understanding	Having both a belief and feeling for <i>A</i> 's emotion (<i>affective- focused empathy</i>)	Having both a belief and feeling for <i>A</i> 's belief (<i>cognitive-focused empathy</i>)	Having a belief and feeling for <i>A</i> 's belief and feeling (<i>full empathy</i>)

The notion of empathic understanding as described above is an important ingredient in the temporal-causal network models for social responses described in Chap. 9, based on Treur (2014), and for joint decision making described in Chap. 10.

7.4 The Emergence of Collective Power

Each individual person can exert a certain amount and direction of power by his or her actions, depending on personal characteristics and states. In a situation where such powers are exerted in different directions by multiple individuals, they can easily annihilate each other, or, metaphorically spoken, result in a kind of Brownian motion where particles move back and forth but do not change place much. In cases that the individual momenta (the individual powers and their directions) have an arbitrary distribution over a population, no serious collective momentum will emerge.

7.4.1 *The Emergence of Collective Action Based on Mirroring*

To obtain emergence of collective power, the individual momenta should converge to a similar direction so that a collective momentum can result. Using another metaphor, this is what happens in the universe when, for example, comets or planets are formed out of smaller particles, based on mutual attraction based on gravitation. More specifically, to obtain collective action within groups, by some mutual tuning process by mirroring, shared person states have to emerge that in an anticipatory sense relate to action, and by which collective power can be developed. Types of internal states relating to action are intentions or preparations. They can be seen as tendencies to perform a specific action; the emergence of shared preparations by mirroring may be quite effective in this sense. However, individual internal processes also play an important role in deciding about actions. In generating actions or behaviours usually options are prepared for which a choice has to be made, and to one of which an ownership or commitment has to be developed. In the recent cognitive and neurological literature much can be found on the mechanisms behind developing ownership of an action (e.g., Voss et al. 2010). In this literature a distinction is made between prior ownership states, among others based on prediction of effects of a prepared action, and retrospective ownership states, for which in addition the monitored execution of the action and the sensed actual effects play an important role; see also Treur (2011, 2012a) or Chap. 8 below. Prior ownership states play an important role in self-control: controlling the actual execution of an action; they also may entail a form of metacognitive functioning to the extent that such states can be made aware and accessible to reflection.

One of the issues that play an important role for both prior and retrospective ownership states, is internal simulation as a means for prediction of the (expected) effects of a prepared action. As discussed earlier, the idea of internal simulation is that in a certain context (which may cover sensed aspects of the external world, but also internal aspects such as the own goals and attitudes) preparation states for actions or bodily changes are activated, which, by prediction links, in turn activate other sensory representation states. The latter states represent the effects of the prepared actions or bodily changes, without actually having executed them. The notion of internal simulation has a longer tradition, and has been put forward, among others, for prediction of effects of one's own prepared motor actions (Becker and Fuchs 1985), imagination (Hesslow 2002), processes in a person's own body related to emotional responding through as-if body loops (Damasio 1994, 1999), and recognition or reading another person's emotions or mind (Goldman 2006).

Thus, based on internal simulation, predictions are made for the effects of prepared actions and based on these, a prior ownership state is generated to support self-control. More specifically, if the predicted effects of a prepared action are valued as satisfactory with respect to the person's goals (prior valuation), this may entail a 'go' decision for the actual execution of the action, thus exerting control over action execution. In contrast, predicted effects valued as less satisfactory for the person's goals may lead to a 'no go' decision for that option. Over the years the idea has developed that retrospective ownership is based on some form of (retrospective valuing of) co-occurrence of predicted effects and sensed actual effects, after execution of the action. This has traditionally been described by a so-called 'comparator model' inspired by cybernetics and control theory (Wolpert 1997). More recently it has been found that to obtain a retrospective ownership state the predicted effect and the sensed actual effect are in fact not compared but added to each other in some integration process (Voss et al. 2010).

Behaviour options usually have emotional responses associated to them relating to a prediction of a rewarding or aversive consequence in the context of the person's goals. Therefore, valuing of options to decide for some behaviour, prior to a choice or in retrospection after a choice was made, have a strong emotional component. In recent neurological literature this has been studied in relation to a notion of value as represented in the amygdala (Bechara et al. 2003; Damasio 1994; Morrison and Salzman 2010); see also Chap. 6. In making a decision for a certain behaviour, experiences with the environment (from the past) play an important role. By a retrospective process, the valuations (and their related emotions) of behaviour options are adapted to the experiences, so that the decision making is adapted to the environment as reflected in these past experiences. In humans parts of the prefrontal cortex (PFC) and other brain areas such as hippocampus, basal ganglia, and hypothalamus have extensive, often bidirectional connections with the amygdala (Morrison and Salzman 2010). Usually emotional responses are triggered by stimuli for which a predictive association is made of a rewarding or aversive consequence, given the context including the person's goals. Feeling these emotions represents a way of experiencing the value of such a prediction, and to which extent it is positive or negative for the person's goals: prior valuation of the option. Similarly, feelings

of satisfaction are an important element of retrospective valuation of what is experienced after behaviour has been chosen.

In emerging collective decision making the individual internal processes have to be dealt with. A mirroring process may help to achieve that a specific preparation option gets a high activation level for all individuals in a group. However, when the own internal processes would keep on driving the person in a different direction, no collectiveness will be achieved. Therefore the mirroring will not only have to address the preparation states, but also the emotion-related valuation states that play a main role in the own internal process. This will be discussed subsequently.

7.4.2 The Role of Feelings and Valuing in the Emergence of Collective Action

In some more detail the situation is as follows. Usually in the individual process of action selection, before a prepared action comes in focus to be executed, an internal simulation to predict the effects of the action takes place: the action is simulated based on prediction links, and in particular for the associated affective effects, based on as-if body loops that predict the body state which is the basis of the related feeling (Damasio 1994, 1999, 2003, 2010). Based on these predicted effects a valuation of the action takes place, which may involve or be mainly based on the associated affective state, as, for example, described in Bechara et al. (2003), Damasio (1994, 1996, 2003), Morrison and Salzman (2010), Murray (2007). The idea here is that by an as-if body loop each option (prepared action) induces a simulated effect including a feeling which is used to value the option. For example, when a negative feeling and value is induced by a particular option, it provides a negative assessment of that option, whereas a positive feeling and value provides a positive assessment. The decision for executing a prepared action is based on the most positive assessment for it.

This simulation process for prepared actions does not only take place for preparations of self-generated actions, but also for intentions or actions from other persons that are observed. In this way by the mirroring process not only a form of action or intention recognition takes place in the form of activation of corresponding own preparation states by mirror neurons, but in addition also the (predicted) effects are simulated, including the affective effects. This provides an emotionally grounded form of understanding of the observed intention or action, including its valuing, which is shared with the observed person; see also Damasio (2010, pp. 102–104).

Given the important role of the feeling states associated to preparations of actions, it may be unrealistic to expect that a common action can be strong when the individual feelings and valuations about such an action have much variation over a group. When only the preparations for options are tuned to each other while in the meantime still the individual internal processes underlying the decision making

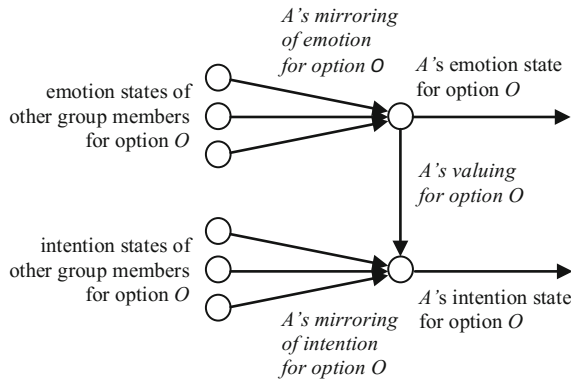


Fig. 7.3 Mirroring processes for both emotions and intentions and their internal interaction

remain a strong drive in a different direction, the overall process may result in no collectiveness at all. To achieve emergence of strong collective action, also a shared feeling and valuation for this action has to develop: also mirroring of the associated emotions has to play an important role. When this is achieved, the collective action has a solid shared emotional grounding: the group members do not only intend to perform that action collectively, but they also share a good feeling about it. In this process social media can play an important facilitating role in that (1) they dramatically strengthen the connections between large numbers of individuals, and (2) they do not only support transfer of, for example, beliefs and intentions as such, but also associated emotions reinforcing them. Thus emergence of collectiveness of action is achieved by not only tuning the preparations or intentions for options to each other, but by also tuning the individual internal processes underlying the decision making for these options; see Fig. 7.3. This double-effective form of contagion enables both the emergence of a collective action and of a solid emotional grounding for this collective action.

The notion of double-effective contagion as described above is an important ingredient in the temporal-causal network model for joint decision making described in Chap. 10.

7.5 Integration of External Effects and Internal Processes

In more realistic cases an interplay occurs between impact of other persons and other internal states. For example, Fig. 7.4 shows an internal model where a certain cognitive state (for example, a sensory representation or belief) has both a cognitive and affective impact on a person's emotions and preparations. Usually such impacts also have feedback loops; an example of this is an as-if body loop (see Sect. 7.2). Therefore, often an internal model consists of a number of cycles, for example, as shown in Fig. 8. In processing, these loops may converge to some equilibrium,

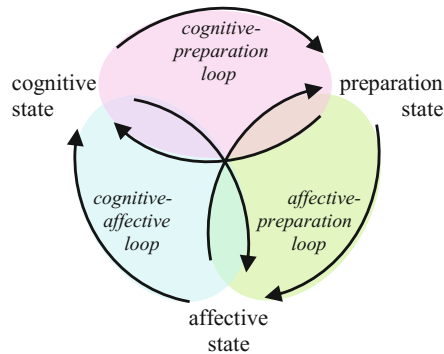


Fig. 7.4 Example of more complex interactions in an internal model

when impact from outside is not changing too fast. Such a way of combination was used in the computational model for emotion-grounded collective decision making described in Hoogendoorn et al. (2011), based on the principles discussed in Sect. 7.4 above. In this case mirroring was applied to both emotion and intention states for any option O :

- *mirroring of emotions* as a mechanism for how emotions felt about a certain considered decision option O in different individuals mutually affect each other
- *mirroring of intentions* as a mechanism for how strengths of intentions (action tendencies) for a certain decision option O in different individuals affect each other.

In the model not only intentions of others, but also a person's emotions affect the person's own intentions (the arrow from affective state to preparation state in Fig. 7.4). In updating the level of an intention state S relating to an option O , the intention states of others for O and the values for the emotion state S' for O were taken into account, and aggregated using the approach indicated above. In simulations in most cases not only a collective decision for an intention was emerging, but also a shared underlying feeling. For more details and simulation results, see Hoogendoorn et al. (2011). Examples of exceptions occur when group members have no openness for others, or are not connected to others.

An example of a more complex computational model is the collective decision making model ASCRIBE addressing an interplay of beliefs, intentions, and emotions, see Fig. 7.5; e.g., Hoogendoorn et al. (2011), Bosse et al. (2013). The internal model used here instantiates part of the general picture of Fig. 7.4. Beliefs instantiate the cognitive, emotions the affective, and intentions the preparation states. In this specific internal model it is assumed that an individual's strength of an intention for a certain decision option depends on the person's beliefs (*cognitive responding*) and emotions (*emotion-related valuing* or *somatic marking*) in relation to that option. Moreover, it is assumed that beliefs may generate certain emotions (*affective responding*), for example of fear, that in turn may affect the strength of beliefs (*affective biasing*). Note that these latter emotion impacts are independent of

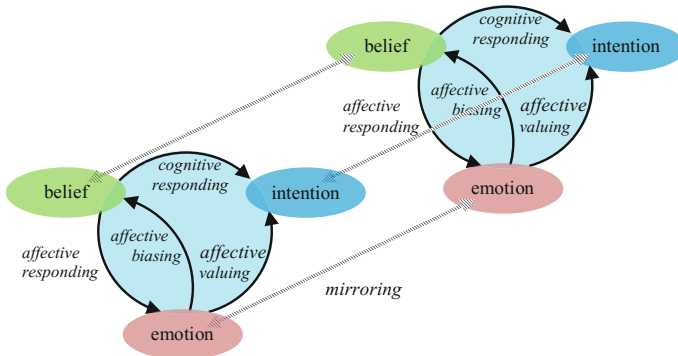


Fig. 7.5 Threefold mirroring integrated with internal interplay of beliefs, emotions and intentions

specific decision options (e.g., a general fear level). Mirroring was used in three different forms (the dotted arrows in Fig. 7.5): of emotions (both fear and emotions felt about a certain decision option O), of beliefs, and of intentions (for a certain decision option O). In the model for the dynamics of intentions, the impact from mirroring is combined with impact from the emotion states and impact from beliefs, in a similar manner as described above. The same applies, for example, to the impact of beliefs on the emotion state. However, in this model also a different type of combination of mirroring and internal processes takes place, involving impact of fear states to beliefs: it is assumed that some of the parameters, for example, for biases and openness with respect to beliefs are affected by fear levels. For more details of this model, including the model specifications for the internal processes of each person and a number of example simulations, see Hoogendoorn et al. (2011); in Bosse et al. (2011, 2013) an application to a real world crowd behaviour case is presented.

7.6 Abstraction of Complex Internal Temporal-Causal Network Models

The temporal-causal network models discussed above were specified as internal models at the cognitive and affective level, and often involve loops between different internal states, for example, loops between cognitive and affective states. However, under certain assumptions such internal models can be abstracted to behavioural temporal-causal network models providing more efficient processing, which is important especially when larger numbers of persons are simulated. In Sharpanskykh and Treur (2010a, b) it is addressed how more complex internal temporal-causal network models can be abstracted to less complex behavioural models.

Temporal-causal network models used for collective social phenomena traditionally are kept simple, and often are specified by simple reactive rules that determine a direct response (output) based on the person's current perception (input). However, in recent years it is more and more acknowledged that in some cases temporal-causal network models specified in the simple format as input–output associations are too limited. Dynamics of internal processes of a person are usually modeled by an internal temporal-causal network model specifying relations between mental states of the person. Often such temporal-causal network models are specified in an executable format following a non-cyclic causal graph. However, for more complex and adaptive types of persons, models may be needed that have a format of dynamical systems including internal loops. Such cyclic interactions are well-known from neurological and brain research areas, for example, loops to model the mutual interaction between affective and cognitive states (see also Figs. 7.5 and 6 above). Thus, although the non-cyclic graph assumption behind most existing models may be useful for the design of (artificial) software agents, it seriously limits applicability for modeling more realistic neurologically inspired processes in a natural context.

In Sharpanskykh and Treur (2010b) an automated transformation is introduced from an internal temporal-causal network model to a behavioural model, abstracting from the internal states. Within this transformation, techniques for loop abstraction are applied by identifying how equilibrium states depend on inputs for these loops. This loop elimination approach can be applied if some underlying assumptions are fulfilled, for example, that the internal dynamics develop an order of magnitude faster than the dynamics external to the person, that the loop indeed reaches an equilibrium, and that the value for this equilibrium can be determined analytically (by solving the equilibrium equations with the input for the loop as parameter). The idea is that when these assumptions are fulfilled, for each received input, before new input arrives, the person computes its internal equilibrium states, and based on that determines its behaviour. More on such abstraction methods can be found in (Sharpanskykh and Treur 2010a, b, 2012).

7.7 Discussion

In this chapter it was discussed how biological mechanisms from the discipline Social Neuroscience can be exploited to obtain biologically grounded temporal-causal network models for social phenomena, covering both cognitive and affective processes, and their interaction. The contents of this chapter are mainly based on Treur (2011b, c, 2012b). Core mechanisms used are mirror neurons and internal simulation. Mirror neurons are certain neurons that are activated due to observation of another person having a corresponding state; e.g., Iacoboni (2008a), Pineda (2009), Rizzolatti and Sinigaglia (2008), Molenberghs et al. (2012), Kilner and Lemon (2013). Internal simulation is internal processing copying an external process, for example another person's mental process; e.g., Damasio (1994, 1999),

Gallese and Goldman (1998), Goldman (2006), Hesslow (1994, 2002, 2012), Barsalou (2009), Marques and Holland (2009), Pezzulo et al. (2013).

It was shown how from a Network-Oriented Modeling perspective, temporal-causal network models can be designed for such processes and used to perform simulation and analysis of the emergence of shared understanding of a group. Furthermore, it was shown how such temporal-causal network models can be used to perform simulation and analysis of the emergence of collective power of a group. This was addressed both in a cognitive or affective or combined sense, so that not only the group members together go for a collective action, but they also share the experience of a good feeling about it, which gives the collective action a solid emotional grounding. It was discussed how such processes depend on the connection strengths between persons, which are strengthened, for example, by social media.

The type of (social) persons modeled from this perspective are integrative in three different manners. In the first place, within the person's individual internal mental processes affective and cognitive processes are not separate mental processes, but are fully integrated in an bidirectional interactive manner: cognitive states lead to associated emotions, and emotions affect cognitive states. Secondly, also preparations and sensory representation states affect each other in a bidirectional interactive manner. Thus, the individual internal functioning is modeled using intensively cyclic processes, instead of the traditional view based on relatively simple non-cyclic processes according to a linear sequence from sensing to preparing for an action. In the third place the modeled social persons integrate their individual internal processes with mirroring processes based on the social context in which they function. As a result temporal-causal network models are obtained that are integrative in multiple respects: integrating cognitive-affective, preparing-sensing, and individual-social impacts. Note that also adaptive processes may take place by which the persons for example change their connections between certain internal states. For such adaptive persons the above elements are also integrated with the adaptation processes. This was left outside the scope of the current chapter, but is addressed in other chapters; for example, Chaps. 2, 5, 6, 11, 12.

The perspective put forward in this chapter has a number of possible application areas. In the first place it can be used to analyse human social processes in groups, crowds or in societies as a whole. The application to crowd behaviour in emergency situations addressed in Bosse et al. (2011, 2013) is an example of such an application. Other cases address, for example, collective decision making, the construction of social reality (Butz 2008), the development of collective consciousness (Combs and Krippner 2008), and global empathy enabling to solve global problems such as climate change (Rifkin 2010), or spiritual and religious experience (Seybold 2010).

A second area of application addresses socio-technological systems that consist of groups that partly consist of humans and partly of devices, such as smartphones, and use of social media. For such mixed groups in addition to analysis of what patterns may emerge, also the design of these devices and media can be an important aim, in order to create a situation that the right types of patterns emerge, for example, with safe evacuation in an emergency situation as a consequence.

A third area of application concerns a close empathic interaction between a human and a device. The importance of computational models in a virtual context for ‘caring’ virtual agents showing empathy has also been well-recognized in the literature; see, for example (Bickmore and Picard 2004). In Sect. 7.3 it has been discussed how such a virtual agent can have empathic understanding by having the same feeling as the human and believing that this is the feeling of the human. This can be shown by a combined communication: showing feeling (nonverbally) and showing knowing (verbally). As a fourth area of application team formation can be addressed. In this area it may be analysed in what way the above perspective provides possibilities that differ compared to already existing approaches.

More variations and details of models as pointed out in this chapter, can be found in the next four chapters. Moreover, references to specific papers can be provided. For example, in Bosse et al. (2012); Laan and Treur (2011) it is shown in more detail how mirroring plays a role in emotion recognition. Examples with both mirroring and control functions can be found in Hendriks and Treur (2010), Treur (2011a, b). In Hendriks and Treur (2010) it is shown how depending on the context through control states, activation of a preparation state has a function in either execution, recognition, imagination or imitation of an action. In Treur (2011a) it is shown how such control states play a role in regulation of different forms of social response patterns, and in Treur (2011b) in prior and retrospective ownership states for an action. Network models for the emergence of shared understanding can be found, for example, in Bosse et al. (2009, 2015) where a model for converging emotion spirals (e.g., of fear) is described. In Hoogendoorn et al. (2011) a model for cognitive states (beliefs), and affective states (fear) with respect to the external world (in mutual relation) is described which shows how for such combined cases shared understanding emerges. Models that have been developed for different types of shared understanding of internal states of persons based on a mirroring mechanism, can be found, for example, in Bosse et al. (2012), Treur et al. (2011a) for affective-focused empathic understanding and social responses, and in Memon and Treur (2012) for full empathic understanding.

References

- W. Barsalou, Simulation, situated conceptualization, and prediction Lawrence. *Phil. Trans. R. Soc. B* **364**, 1281–1289 (2009)
- A. Bechara, H. Damasio, A.R. Damasio, Role of the amygdala in decision-making. *Ann. N.Y. Acad. Sci.* **985**, 356–369 (2003)
- W. Becker, A.F. Fuchs, Prediction in the oculomotor system: smooth pursuit during transient disappearance of a visual target. *Exp. Brain Res.* **57**, 562–575 (1985)
- T.W. Bickmore, R.W. Picard, Towards caring machines, in *Proceedings of CHI'04*, ed. by E. Dykstra-Erickson, M. Tscheligi (ACM, 2004), pp. 1489–1492
- T. Bosse, R. Duell, Z.A. Memon, J. Treur, C.N. Wal van der, A multi-agent model for emotion contagion spirals integrated within a supporting ambient agent model, in *Proceedings PRIMA'09. LNAI*, vol. 5925, ed. by J.-J. Yang et al. (Springer, 2009), pp. 48–67

- T. Bosse, M. Hoogendoorn, M.C.A. Klein, J. Treur, C.N. Wal van der, Agent-based analysis of patterns in crowd behaviour involving contagion of mental states, in *Proceedings of IEA/AIE'11, Part II. LNAI*, vol. 6704, ed. by K.G. Mehrotra et al. (Springer, 2011), pp. 566–577
- T. Bosse, Z.A. Memon, J. Treur, A cognitive and neural model for adaptive emotion reading by mirroring preparation states and hebbian learning. *Cogn. Syst. Res.* **12**, 39–58 (2012a). doi [10.1016/j.cogsys.2010.10.003](https://doi.org/10.1016/j.cogsys.2010.10.003)
- T. Bosse, Z.A. Memon, J. Treur, A cognitive and neural model for adaptive emotion reading by mirroring preparation states and hebbian learning. *Cogn. Syst. Res.* **12**, 39–58 (2012b)
- T. Bosse, M. Hoogendoorn, M.C.A. Klein, J. Treur, C.N. van der Wal, A. van Wissen, Modelling collective decision making in groups and crowds: integrating social contagion and interacting emotions, beliefs and intentions. *Auton. Agent. Multi-Agent Syst. J.* **27**, 52–84 (2013)
- T. Bosse, R. Duell, Z.A. Memon, J. Treur, C.N. van der Wal, Agent-based modelling of emotion contagion in groups. *Cog. Comput. J.* **7**, 111–136 (2015)
- M. Brass, S. Spengler, The inhibition of imitative behaviour and attribution of mental states, in *Social Cognition: Development, Neuroscience, and Autism*, ed. by T. Striano, V. Reid (Wiley-Blackwell, 2009), pp. 52–66
- M.V. Butz, Intentions and mirror neurons: from the individual to overall social reality. *Constr. Found.* **3**, 87–89 (2008)
- J.T. Cacioppo, G.G. Berntson, *Social Neuroscience* (Psychology Press, 2005)
- J.T. Cacioppo, P.S. Visser, C.L. Pickett, *Social Neuroscience: People Thinking About Thinking People* (MIT Press, Cambridge, MA, 2006)
- S. Cochin, B. Barthelemy, S. Roux, J. Martineau, Observation and execution of movement similarities demonstrated by quantified electroencephalography. *Euro. J. Neurosci.* **11**, 1839–1842 (1999)
- A. Combs, S. Krippner, Collective consciousness and the social brain. *J. Consci. Stud.* **15**, 264–276 (2008)
- A. Damasio, *Descartes' Error: Emotion, Reason and the Human Brain* (Papermac, London, 1994)
- A. Damasio, The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophic. Trans. R. Soc. Biol. Sci.* **351**, 1413–1420 (1996)
- A. Damasio, *The Feeling of What Happens. Body and Emotion in the Making of Consciousness* (Harcourt Brace, New York, 1999)
- A. Damasio, *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain* (Vintage books, London, 2003)
- A.R. Damasio, *Self Comes to Mind: Constructing the Conscious Brain* (Pantheon Books, NY, 2010)
- F. De Vignemont, T. Singer, The empathic brain: how, when and why? *Trends Cogn. Sci.* **10**, 437–443 (2006)
- J. Decety, J.T. Cacioppo (eds.), *Handbook of Social Neuroscience* (Oxford University Press, 2010)
- J. Decety, W. Ickes, *The Social Neuroscience of Empathy* (MIT Press, 2009)
- I. Fried, R. Mukamel, G. Kreiman, Internally generated preactivation of single neurons in human medial frontal cortex predicts volition. *Neuron* **69**, 548–562 (2011)
- V. Gallese, The roots of empathy: the shared manifold hypothesis and the neural basis of intersubjectivity. *Psychopathology* **36**, 171–180 (2003)
- V. Gallese, A. Goldman, Mirror neurons and the simulation theory of mind-reading. *Trend Cogn. Sci.* **2**, 493–501 (1998)
- V. Gallese, L. Fadiga, L. Fogassi, G. Rizzolatti, Action recognition in the premotor cortex. *Brain* **119**, 593–609 (1996)
- H.J. Gastout, J. Bert, EEG changes during cinematographic presentation. *Electroencephalograp. Clin. Neurophysiol.* **6**, 433–444 (1954)
- A.I. Goldman, *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading* (Oxford Univ. Press, New York, 2006), p. 2006

- S.T. Grafton, M.A. Arbib, L. Fadiga, G. Rizzolatti, Localisation of grasp representations in humans by PET: 2. Observation compared with imagination. *Exp. Brain Res.* **112**, 103–111 (1996)
- E. Harmon-Jones, P. Winkielman (eds.), *Social Neuroscience: Integrating Biological and Psychological Explanations of Social Behavior* (Guilford, New York, 2007)
- M. Hendriks, J. Treur, Modeling super mirroring functionality in action execution, imagination, mirroring, and imitation, in *Proceedings of ICCCI'10, Part I LNAI*, vol. 6421, ed. by J.-S. Pan et al. (Springer, 2010), pp. 330–342
- G. Hesselow, Will neuroscience explain consciousness? *J. Theoret. Biol.* **171**, 29–39 (1994)
- G. Hesselow, Conscious thought as simulation of behaviour and perception. *Trends Cogn. Sci.* **6**, 242–247 (2002)
- G. Hesselow, The current status of the simulation theory of cognition. *Brain Res.* **1428**, 71–79 (2012)
- G. Hickok, Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *J. Cogn. Neurosci.* **21**, 1229–1243 (2009)
- M. Hoogendoorn, J. Treur, C.N. van der Wal, A. van Wissen, Agent-based modelling of the emergence of collective states based on contagion of individual states in groups. *Trans. Comput. Collect. Intell.* **3**, 152–179 (2011)
- M. Iacoboni, *Mirroring People: The New Science of How We Connect with Others* (Farrar, Straus & Giroux, New York, 2008a)
- M. Iacoboni, Mesial frontal cortex and super mirror neurons. *Beh. Brain Sci.* **31**, 30–30 (2008b)
- M. Iacoboni, I. Molnar-Szakacs, V. Gallese, G. Buccino, J.C. Mazziotta, G. Rizzolatti, Grasping the intentions of others with one's own mirror neuron system. *PLoS Biol.* **3**, e79 (2005)
- W. James, What is an emotion. *Mind* **9**, 188–205 (1884)
- C. Keysers, V. Gazzola, Social neuroscience: mirror neurons recorded in humans. *Curr. Biol.* **20**, 253–254 (2010)
- J.M. Kilner, R.N. Lemon, What we know currently about mirror neurons. *Curr. Biol.* **23**, R1057–R1062 (2013)
- Y. van der Laan, J. Treur, An agent model for computational analysis of mirroring dysfunctioning in autism spectrum disorders, in *Proceedings of IEA/AIE'11, Part I. LNAI*, vol. 6703, ed. by K. G. Mehrotra et al. (Springer, 2011), pp. 306–316
- T. Lipps, Einfühlung, innere Nachahmung und Organempfindung. *Archiv für die gesamte Psychologie* **1**, 465–519 (1903)
- H.G. Marques, O. Holland, Architectures for functional imagination. *Neurocomputing* **72**, 743–759 (2009)
- Z.A. Memon, J. Treur, An agent model for cognitive and affective empathic understanding of other agents. *Trans. Comput. Collect. Intell.* **6**, 56–83 (2012). Earlier, shorter version in *Proceedings of ICCCI'09, Part A LNAI*, vol. 5796, ed. by N.T. Nguyen, R. Kowalczyk, S.M. Chen. (Springer, 2009), pp. 279–293
- P. Molenberghs, R. Cunnington, J.B. Mattingley, Brain regions with mirror properties: a meta-analysis of 125 human fMRI studies. *Neurosci. Biobehav. Rev.* **36**, 341–349 (2012)
- S.E. Morrison, C.D. Salzman, Re-valuing the amygdala. *Curr. Opin. Neurobiol.* **20**, 221–230 (2010)
- R. Mukamel, A.D. Ekstrom, J. Kaplan, M. Iacoboni, I. Fried, Single-neuron responses in humans during execution and observation of actions. *Curr. Biol.* **20**, 750–756 (2010)
- E.A. Murray, The amygdala, reward and emotion. *Trends Cogn. Sci.* **11**, 489–497 (2007)
- G. Pezzulo, M. Candidi, H. Dindo, L. Barca, Action simulation in the human brain: twelve questions. *New Ideas Psychol* **31**, 270–290 (2013)
- J.A. Pineda (ed.), *Mirror Neuron Systems: the Role of Mirroring Processes in Social Cognition* (Humana Press Inc., 2009)
- S.D. Preston, F.B.M. Waal de, Empathy: its ultimate and proximate bases. *Behav. Brain Sci.* **25**, 1–72 (2002)
- J. Rifkin, *The Empathic Civilization: The Race to Global Consciousness in a World in Crisis* (Tarcher Penguin, 2010)

- G. Rizzolatti, L. Craighero, The mirror neuron system. *Ann. Rev. Neurosci.* **27**, 169–192 (2004)
- G. Rizzolatti, C. Sinigaglia, *Mirrors in the Brain: How Our Minds Share Actions and Emotions* (Oxford University Press, 2008)
- G. Rizzolatti, L. Fadiga, V. Gallese, L. Fogassi, Premotor cortex and the recognition of motor actions. *Cogn. Brain Res.* **3**, 131–141 (1996a)
- G. Rizzolatti, L. Fogassi, M. Matelli, et al., Localisation of grasp representations in humans by PET: 1. observation and execution. *Exp. Brain Res.* **111**, 246–252 (1996b)
- K.S. Seybold, Biology of spirituality. *Perspect. Sci. Christ. Faith* **62**, 89–98 (2010)
- S.G. Shamay-Tsoory, Empathic processing: its cognitive and affective dimensions and neuroanatomical basis, in *The Social Neuroscience of Empathy*, ed. by J. Decety, W. Ickes (Cambridge, Mass, MIT Press, 2008), pp. 215–232
- S.G. Shamay-Tsoory, The neural bases for empathy. *Neuroscience* **17**, 18–24 (2011)
- A. Sharpanskykh, J. Treur, Abstraction relations between internal and behavioural agent models for collective decision making, in *Proceedings of ICCCI'10. LNAI*, vol. 6421, ed. by J.-S. Pan et al. (Springer, 2010a), pp. 39–53
- A. Sharpanskykh, J. Treur, Behavioural abstraction of agent models addressing mutual interaction of cognitive and affective processes, in *Proceedings of BI'10. LNAI*, vol. 6334, ed. by Y. Yao et al. (Springer, 2010b), pp. 67–77
- A. Sharpanskykh, J. Treur, Abstraction relations between internal and behavioural agent models for collective decision making. *Web Intell. Agent Syst. J.* **10**, 465–484 (2012)
- T. Singer, S. Leiberg, Sharing the emotions of others: the neural bases of empathy, in *The Cognitive Neurosciences*, 4th ed, ed. by M.S. Gazzaniga (MIT Press, 2009), pp. 973–986
- J. Treur, A cognitive agent model displaying and regulating different social response patterns, in *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence, IJCAI'11*, ed. by T. Walsh (AAAI Press, 2011a), pp. 1735–1742. doi:[10.5591/978-1-57735-516-8/IJCAI11-291](https://doi.org/10.5591/978-1-57735-516-8/IJCAI11-291)
- J. Treur, A cognitive agent model incorporating prior and retrospective ownership states for actions, in *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence, IJCAI'11*, ed. by T. Walsh. (AAAI Press, 2011b), pp. 1743–1749. doi [10.5591/978-1-57735-516-8/IJCAI11-292](https://doi.org/10.5591/978-1-57735-516-8/IJCAI11-292)
- J. Treur, From mirroring to the emergence of shared understanding and collective power, in *Proceedings of ICCCI'11, Part I. Lecture Notes in Artificial Intelligence*, vol. 6922, ed. by P. Jedrejowicz, N.T. Nguyen, K. Hoang (Springer, Heidelberg, 2011c), pp. 1–16
- J. Treur, A computational agent model incorporating prior and retrospective ownership states for actions. *Biol. Insp. Cogn. Architect. J.* **2**, 54–67 (2012a)
- J. Treur, Biological and computational perspectives on the emergence of social phenomena: shared understanding and collective power. *Trans. Comput. Collect. Intell.* **8**, 168–191 (2012b)
- J. Treur, Displaying and regulating different social response patterns: a computational agent model. *Cogn. Comput. J.* **6**, 182–199 (2014)
- M. Voss, J. Moore, M. Hauser, J. Gallinat, A. Heinz, P. Haggard, Altered awareness of action in schizophrenia: a specific deficit in predicting action consequences. *Brain* **133**, 3104–3112 (2010)
- D.M. Wolpert, Computational approaches to motor control. *Trends Cogn. Sci.* **1**, 209–216 (1997)

Chapter 8

Am I Going to Do This? Is It Me Who Did This?

Prior and Retrospective Ownership States for Actions

Abstract In this chapter a Network-Oriented Modeling approach is used to obtain a temporal-causal network model that generates prior and retrospective ownership states for an action based on principles from recent neurological theories. A prior ownership state is affected by prediction of the effects of a prepared action, and exerts control by strengthening or suppressing actual execution of the action. A retrospective ownership state depends on whether the sensed consequences of an executed action co-occur with the predicted consequences, and is the basis for acknowledging authorship of actions, for example, in a social context. It is shown how a number of known phenomena can occur. For example, scenarios are shown for vetoing a prepared action due to unsatisfactory predicted effects, and for mirroring an observed action performed by another person, without imitating the action. Moreover, it is shown how poor action effect prediction capabilities can lead to reduced retrospective ownership states, as in persons suffering from Schizophrenia. The obtained temporal-causal network model can be used as a basis for simulation-based training, for example, to develop a virtual patient based on the model so that a psychiatrist or psycho-therapist (e.g., during his or her education) can gain insight in the processes in certain types of patients, or to analyse how effective a certain form of therapy can be. A second type of application is in the area of gaming or virtual stories in which, for example, persons with deviations in ownership states play a role and based on that show unexpected behaviour.

8.1 Introduction

In the cognitive and neurological literature the notion of ownership of an action has received much attention: in how far does a person attribute an action to him or herself, or to another person. For example, persons suffering from schizophrenia may easily attribute self-generated actions to (real or imaginary) other persons. One of the issues that plays an important role both in the execution decisions for an action, and in its attribution, is the prediction of the (expected) effects of the action,

based on internal simulation starting from the preparation of the action (e.g., Wolpert 1997; Haggard 2008). If these predicted effects are satisfactory, this may entail a ‘go’ decision for the execution of the action, thus exerting control over action execution. In contrast, less satisfactory predicted effects may lead to vetoing a prepared action: a ‘no go’ decision.

Predicted action effects play an important role in attribution of the action to a person after it has been performed. In neurological research it has been found that poor predictive capabilities are a basis for false attributions of actions, for example, for patients suffering from schizophrenia; (e.g., Synofzik et al. 2010; Voss et al. 2010). The traditional approach is that comparison of predicted sensory effects and sensed actual effects (after execution of the action) is an important condition for proper retrospective self-attribution of a self-generated action. The so-called ‘comparator model’ (e.g., Feinberg 1978; Frith 1992; Wolpert 1997; Frith et al. 2000) is a computational formalisation of this idea. In contrast to this, it has been reported more recently that the predicted sensory effect and the sensed actual effect are not simply compared, but integrated with each other (e.g., Moore and Haggard 2008; Synofzik et al. 2010; Voss et al. 2010).

The temporal-causal network model presented in this chapter aims at providing a computational formalisation for the perspective reported, for example, in the latter recent literature. In designing this computational network model, in line with this literature a distinction was made between prior ownership states, among others based on prediction of sensory effects of a prepared action, and retrospective ownership states, for which in addition the monitored execution of the action and the sensed actual effects are used. Within the network model prior ownership states play an important role in controlling the actual execution of actions (go/no-go decisions, vetoing), whereas retrospective ownership states are important for acknowledging authorship of an action in a social context, but also may play a role in reflection on one’s own functioning and personal learning and development (e.g., learning from less optimal choices).

The temporal-causal network model can be used as a basis for development of a virtual patient that can be used by therapists, for example during their education to increase their insights in the dynamics of processes involving ownership, and deviations that may occur in these processes. Moreover, in the area of gaming or virtual stories the model can be used to develop virtual characters that show special behaviour in relation to certain characteristics concerning ownership.

In this chapter, in Sect. 8.3 the temporal-causal network model is presented, based on neurological principles discussed in Sect. 8.2. Section 8.4 illustrates the model by presenting four different scenarios, among which vetoing an action because of unsatisfactory predicted effects, execution of an action with false attribution due to poor predictive capabilities as happens in schizophrenia, and mirroring an observed action of another person and properly attributing it to the other person. In Sect. 8.5 it is discussed how elements of the model can be related to recent neurological findings. Finally, Sect. 8.6 is a discussion.

8.2 Neurological Background

Within the literature over the years the idea developed that an important condition for proper retrospective self-attribution of a self-generated action is co-occurrence of predicted sensory effects and sensed actual effects (after execution of the action); e.g., Moore and Haggard (2008). In what is called the *comparator model* or the *central monitoring theory* (e.g., Feinberg 1978; Wolpert 1997; Frith 1992; Frith et al. 2000; David et al. 2008), the predicted effect and the sensed actual effect are compared or matched. In Fig. 8.1 a schematic overview of the comparator model is shown, comparable to similar pictures shown in Wolpert (1997) and David et al. (2008).

The notion of comparison was originally taken from the area of cybernetics, and assumes that two activation strengths are compared. For example, if a binary case is assumed, two signals 1 provide an outcome 0 (i.e., no difference: self agency), and a combination of 0 and 1 gives an outcome 1 (no self agency). Also two strengths 0 would give a comparison outcome 0 (self agency) in this model. Moreover, for the nonbinary case, for example, two equal values 0.3 (or 0.5) would give an outcome 0 as well, which is the same as the outcome for two values 1 (self agency). Such simple examples, especially for lower but equal values may not always be considered intuitive.

In recent years it has been reported in neurological literature that the predicted effect and the sensed actual effect are not simply compared or matched, as claimed in the comparator model, but in fact are added to each other in some integration process (e.g., Moore and Haggard 2008; Synofzik et al. 2010; Voss et al. 2010). In such a case the outcome would not always be the same as for the comparator model. For example, integrating two equal levels 1 will provide an outcome (self agency) that is quite different from the outcome for two equal lower values 0.3 or two equal values 0. Moreover, the integration process involves a temporal element which is not taken into account in the comparator model: first the predicted sensory effect starts to become active whereas later on the sensing of the actual effect in some sense is added to it.

Furthermore, another peculiar aspect that plays a role here is that within the process, based on the predicted sensory effect, suppression of the sensed effect takes place, which has been reported since a longer time, for example, around the question ‘Why can’t you tickle yourself?’; e.g., Weiskrantz et al. (1971), Claxton

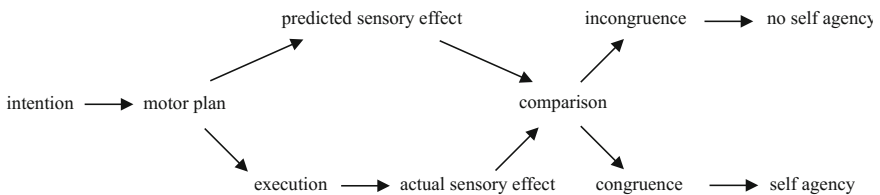


Fig. 8.1 Overview of the comparator model

(1975), Collins et al. (1998), Blakemore et al. (1999, 2000a, b), Fournier et al. (2002) and Blakemore and Frith (2003). This shows an interesting phenomenon. On the one hand the predicted effect suppresses the sensed effect, but on the other hand the two effects are combined with each other to obtain a basis for a (retrospective) proper attribution of the action. From a logical point of view these findings from the neurological literature may sound paradoxal, but from a dynamical systems perspective this may well be modeled adequately in a temporal manner, thus providing a more realistic type of model than would be possible in a purely logical setting.

Another element, put forward in Moore and Haggard (2008), is the distinction between ownership based on prediction (prior to execution), and ownership based on inference after execution of the action (in retrospect):

Our results suggest that both predictive and inferential processes contribute to the conscious awareness of operant action. The relative contribution of each of these processes seems to be context dependent. When we can predict the consequences of our actions, as in a high action-effect contingency block, the awareness of action reflects these predictions. This would provide us with a predictive sense of our own agency. In addition, our results show clear evidence that inferential processes also influence the conscious awareness of operant action (Moore and Haggard 2008 p. 142).

Also here it is put forward that the temporal pattern may be interesting to take into account:

The interaction between predictive and inferential processes is of particular interest. (...) The time course over which information about action is built up may be an important clue to this interaction. Specifically, predictive information about actions and effects may operate only over the brief timescale of motor preparation, and may be discarded when sensory evidence makes inference possible. Predictions may be based on an internal forward model (Blakemore et al. 2002). This representation is available for the control of action and on-line guidance of behaviour, but does not outlast the current action. Sensory feedback provides more precise evidence about actions and their effects. This evidence becomes available only after a short sensory delay, but can then be transferred to memory. Thus, reliable and enduring sensory evidence replaces short-lived predictive estimates. We suggest that awareness of action therefore switches from a predictive to an inferential source as the action itself occurs, and as sensory information becomes available. This time-varying mixture of predictive and inferential information may ensure that our experience of our own action is an optimal reflection of the actual relation between our voluntary motor system and the outside world (Moore and Haggard 2008, pp. 142–143).

The issues and perspectives briefly reviewed above have been used as a basis for the temporal-causal network model presented below. More specifically, the following have been taken as a point of departure:

- (1) action effect *prediction* from preparation of an action *a* to sensory representation of effect *b*
- (2) *suppressing* the sensory representation of effect *e* after this effect was predicted and action *a* was initiated
- (3) a *prior ownership* state depends on preparation for the action, predicted effects, and context

- (4) a *retrospective ownership* state depends on a combination of predicted sensory effects of the action and action effects sensed afterwards (a form of integration)
- (5) a prior ownership state exerts *control over the execution* of a prepared action (go/no-go decision, vetoing)
- (6) a retrospective ownership state is an *internal state* that also can lead to *acknowledging authorship* of the action, for example, in social context.

In Sect. 8.5 it will be discussed in some more detail how the concepts used in the model relate to neurological concepts.

8.3 A Temporal-Causal Network Model for Ownership

The temporal-causal network model presented below has been designed with the issues discussed in Sect. 8.2 as a point of departure [in particular (1)–(6)].

8.3.1 Conceptual Representation of the Temporal-Causal Network Model

For a graphical conceptual representation of the model, see Fig. 8.2; here the circles denote the states explained in Table 8.1.

In this model, s denotes a stimulus, c a context, a an action, and e a world state affected by the action. Examples of contexts are another person B which is observed, or the person *self*. The effect state e is considered to be positive for the person (e.g., in accordance with a goal). Note that these are used as parameters in the names of states so that a structured naming convention is obtained; they are not values. The states used in the model are summarised in Table 8.1. Sensor states ss_w

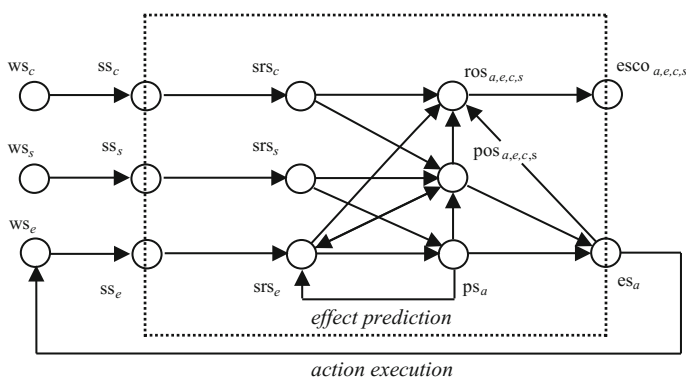


Fig. 8.2 Graphical conceptual representation of the temporal-causal network model

Table 8.1 State properties used

Notation	Description
s	Stimulus
a	Action
e	Action effect
c	Context (person self, or another person B , or another context)
ws_W	World state for W (W is a context c , stimulus s , or effect e)
ss_W	Sensor state for W (W is a context c , stimulus s , or effect e)
srs_W	Sensory representation state for W (W is a context c , stimulus s , or effect e)
ps_a	Preparation state for action a
es_a	Execution state for action a
$pos_{a,e,c,s}$	Prior ownership state for action a with effect e , context c , and stimulus s
$ros_{a,e,c,s}$	Retrospective ownership state for action a with effect e , context c , and stimulus s
$esc_{a,e,c,s}$	Execution state for communication of ownership of action a with effect e , context c , and stimulus s

for a world property W are assumed to be generated (i.e., get nonzero activation values) by sensing or observing the world. This includes the context c which can be *self* or another person B . As expressed in (3) and (4) in Sect. 8.2, the temporal-causal network model distinguishes prior and retrospective ownership states for actions, indicated by $pos_{a,e,c,s}$ and $ros_{a,e,c,s}$, respectively (see Fig. 8.2). These states are taken specific for a given action a , effect e , context c , and stimulus s (triggering preparation of a). When the context c is *self*, an ownership state for c indicates self-ownership attribution, whereas for context c an observed person B , it indicates ownership attributed to B . Note that the stimulus s triggering preparation of action a can be of any type; for social scenarios, it can be taken as a body state (e.g., face expression) of the other person B which is observed. An action effect state e can be any state of the world (possibly including body states).

In accordance with (3) in Sect. 8.2, the prior ownership state $pos_{a,e,c,s}$ is affected by the preparation state ps_a for the action a , the sensory representation srs_e of the (predicted) effect e , the sensory representation srs_s of the stimulus s , and the sensory representation srs_c of the context c ; see the four arrows to $pos_{a,e,c,s}$ in Fig. 8.2. Similarly, as expressed in (4) in Sect. 8.2, the retrospective ownership state $ros_{a,e,c,s}$ is affected by the sensory representation srs_c of the context c , the sensory representation srs_e of the effect e of the action, the prior ownership state $pos_{a,e,c,s}$, and the execution es_a of the action a ; see the arrows to $ros_{a,e,c,s}$ in Fig. 8.2.

Action prediction, expressed in (1) in Sect. 8.2, is modeled by the connection from the action preparation ps_a to the sensory representation srs_e of the effect e . Suppression of the sensory representation of the effect, expressed as (2) in Sect. 8.2, is modeled by the (inhibiting) connection from the prior ownership state $pos_{a,e,c,s}$ to sensory representation srs_e . The control exerted by the prior ownership state, expressed in (5) in Sect. 8.2, is modeled by the connection from $pos_{a,e,c,s}$ to es_a . Finally, acknowledging of ownership, expressed in (6) in Sect. 8.2, is modeled

by the connection from the retrospective ownership state $ros_{a,e,c,s}$ to the communication execution state $esc_{a,e,c,s}$.

Connections between states (the arrows in Fig. 8.2) have weights ω_k , as indicated in Table 8.2. In this table the column LP refers to the (temporally) Local Properties LP1 to LP9 presented below. A weight ω_k has a value between -1 and 1 and may depend on the specific context c , stimulus s , action a and/or effect state e involved. By varying these connection strengths, different possibilities for the repertoire offered by the model can be realised. Note that usually weights are assumed non-negative, except for the inhibiting connections, such as ω_{2o} (see Table 8.2) which models suppression of the sensory representation of effect e .

8.3.2 Numerical Representation of the Temporal-Causal Network Model

Below, the dynamics following the connections between the states in Fig. 8.2 are described in the form of a numerical representation of the model. This is done for each state by a dynamic property specifying how the activation value for this state is updated based on the activation values of the states connected to it (the incoming arrows in Fig. 8.2). The model is based on a Network-Oriented Modeling approach (see Chap. 1), in particular, the temporal-causal network modeling approach described in Chap. 2. Note that modeling the causal relations discussed in neurological literature and shown in Fig. 8.2 does not follow specific single neurons but

Table 8.2 Overview of the connections and their weights

From states	To state	Weight	Connection type
ss_W	srs_W	ω_1	Representing
ps_a	srs_e	ω_2	Prediction
$pos_{a,e,self,s}$		ω_{2o}	Suppressing
ss_b		ω_3	Representing
srs_s	ps_a	ω_4	Responding to s
srs_e		ω_5	Amplifying
srs_c	$pos_{a,e,c,s}$	ω_6	Prior owning
srs_s		ω_7	
srs_e		ω_8	
ps_a		ω_9	
$pos_{a,e,self,s}$	es_a	ω_{10}	Controlling
ps_a		ω_{11}	Executing action
es_a	ws_e	ω_{12}	Effectuating
ws_W	ss_W	ω_{13}	Sensing
srs_c	$ros_{a,e,c,s}$	ω_{14}	Retrospective owning
srs_b		ω_{15}	
$pos_{a,e,c,s}$		ω_{16}	
es_a		ω_{17}	
$ros_{a,e,c,s}$	$esc_{a,e,c,s}$	ω_{18}	Executing communication

abstracts them to cognitive or mental states. By this abstraction neurological knowledge is lifted to a mental (cognitive/affective) modeling level. See also the discussion in Bickle (1998) which he illustrates for the higher level (e.g., folk psychological) in relation to the lower-level (e.g., neurobiological) explanation in the context of Hawkin and Kandel's (1984a, b) work; see Chap. 2, Sect. 2.3.

Activation levels of states affect activation levels of other states (to which they are causally connected; see the arrows in Fig. 8.2). In numerical representation these are specified as Local Properties (LP) for each of the different states. Such a specification uses a parameter η as a speed factor, indicating the speed by which the activation level is updated upon received input from other states. For a state causally depending on multiple other states, values for incoming activation levels are combined, using a combination function $c(\dots)$ for which different choices can be made.

In the example simulations, for the states that are affected by only one state (i.e., in LP1, LP6, LP7, LP9), $c(\dots)$ is chosen the identity function $c(V) = \text{id}(V) = V$, and for the other states $c(\dots)$ is chosen the advanced logistic sum function

$$\mathbf{alogistic}_{\sigma, \tau}(V_1, \dots, V_k) = \left(\frac{1}{1 + e^{-\sigma(V_1 + \dots + V_k - \tau)}} - \frac{1}{1 + e^{\sigma\tau}} \right) (1 + e^{-\sigma\tau})$$

Other types of combination functions might be used as well.

In the text below, for the numerical representation semiformal descriptions are given, together with a formal specification in differential equation format (one differential equation per Local Property LP), summarized in Box 8.1.

The first property LP1 describes how sensory representations are generated for context c and stimulus s , and effect state e (together indicated by variable W).

LP1 Sensory representation for a sensor state

If the sensor state for W has level X_1

and the sensory representation of W has level X_2

then after duration Δt the sensory representation of W will have level $X_2 + \eta [c(\omega_1 X_1) - X_2] \Delta t$.

$$\mathbf{dsrs}_W(t)/\mathbf{dt} = \eta [c(\omega_1 \text{ss}_W(t)) - \text{srs}_W(t)]$$

$$\text{srs}_W(t + \Delta t) = \text{srs}_W(t) + \eta [c(\omega_1 \text{ss}_W(t)) - \text{srs}_W(t)] \Delta t$$

The sensory representation of an effect state e as described by property LP2 is not only affected by a corresponding sensor state for e (which in turn is affected by the world state), as in LP1, but also by two action-related states:

- via the predictive loop by a preparation state, to predict the effect e of a prepared action a (see (1) in Sect. 8.2)
- by an inhibiting connection from the prior self-ownership state, to suppress the sensory representation of the effect e of the action a , once it is initiated [see (2) in Sect. 8.2].

This is expressed in dynamic property LP2. Note that for this suppressing effect the connection weight ω_{2o} from prior ownership state for action a to sensory representation for effect e is chosen negative, for example $\omega_{2o} = -1$.

LP2 Sensory representation for an effect state

If the preparation state for action a has level X_1
 and the prior self-ownership of action a for e , self, and s has level X_2
 and the sensor state for state e has level X_3
 and the sensory representation of state e has level X_4
 then after duration Δt the sensory representation of state e will have level $X_4 + \eta [c(\omega_2 X_1, \omega_{2o} X_2, \omega_3 X_3) - X_4] \Delta t$.

$$\frac{dsrs_e(t)}{dt} = \eta [c(\omega_2 ps_a(t), \omega_{2o} pos_{a,e,self,s}(t), \omega_3 ss_e(t)) - srs_e(t)]$$

$$srs_e(t + \Delta t) = srs_e(t) + \eta [c(\omega_2 ps_a(t), \omega_{2o} pos_{a,e,self,s}(t), \omega_3 ss_e(t)) - srs_e(t)] \Delta t$$

 Preparation for action a is affected by a sensory representation of stimulus s (triggering the action), and also strengthened by predicted effect e of the action:

LP3 Preparing for an action

If sensory representation of s has level X_1
 and sensory representation of e has level X_2
 and the preparation for action a has level X_3
 then after duration Δt the preparation state for action a will have level $X_3 + \eta [c(\omega_4 X_1, \omega_5 X_2) - X_3] \Delta t$.

$$\frac{dps_a(t)}{dt} = \eta [c(\omega_4 srs_s(t), \omega_5 srs_e(t)) - ps_a(t)]$$

$$ps_a(t + \Delta t) = ps_a(t) + \eta [c(\omega_4 srs_s(t), \omega_5 srs_e(t)) - ps_a(t)] \Delta t$$

 Prior ownership of an action a is generated by LP4 [see (3) in Sect. 8.2].

LP4 Generating a prior ownership state

If the sensory representation of context c has level X_1
 and the sensory representation of s has level X_2
 and sensory representation of e has level X_3
 and the preparation for action a has level X_4
 and prior ownership of a for e , c , and s has level X_5
 then after Δt prior ownership of a for c , s , and e will have level $X_5 + \eta [c(\omega_6 X_1, \omega_7 X_2, \omega_8 X_3, \omega_9 X_4) - X_5] \Delta t$.

$$\frac{dpos_{a,e,c,s}(t)}{dt} = \eta [c(\omega_6 srs_c(t), \omega_7 srs_s(t), \omega_8 srs_e(t), \omega_9 ps_a(t)) - pos_{a,e,c,s}(t)]$$

$$pos_{a,e,c,s}(t + \Delta t) = pos_{a,e,c,s}(t) + \eta [c(\omega_6 srs_c(t), \omega_7 srs_s(t), \omega_8 srs_e(t), \omega_9 ps_a(t)) - pos_{a,e,c,s}(t)] \Delta t$$

In case the context c is self, the prior ownership state strengthens the initiative to perform a as a self-generated action: executing a prepared action depends on whether a prior self-ownership state (for the person self) is available for this action (see (5) in Sect. 8.2). This models control over the actual execution of the action

(go/no-go decision) and can, for example, be used to veto the action in a late stage of preparation. This is modeled by LP5.

LP5 Action execution

If prior ownership of a for e , self, and s has level X_1

and preparation for action a has level X_2

and the action execution state for a has level X_3

then after duration Δt the action execution state for a will have

level $X_3 + \eta [c(\omega_{10}X_1, \omega_{11}X_2) - X_3] \Delta t$.

$$\text{des}_a(t)/dt = \eta [c(\omega_{10}\text{pos}_{a,e,\text{self},s}(t), \omega_{11}\text{ps}_a(t)) - \text{es}_a(t)]$$

$$\text{es}_a(t + \Delta t) = \text{es}_a(t) + \eta [c(\omega_{10}\text{pos}_{a,e,\text{self},s}(t), \omega_{11}\text{ps}_a(t)) - \text{es}_a(t)] \Delta t$$

Property LP6 describes in a straightforward manner how execution of action a affects the world state e .

LP6 From action execution to effect state

If the execution state for action a has level X_1 ,

and world state e has level X_2

then after Δt world state e will have level $X_2 + \eta [c(\omega_{12}X_1) - X_2] \Delta t$.

$$\text{dws}_b(t)/dt = \eta [c(\omega_{12}\text{es}_a(t)) - \text{ws}_e(t)]$$

$$\text{ws}_b(t + \Delta t) = \text{ws}_b(t) + \eta [c(\omega_{12}\text{es}_a(t)) - \text{ws}_e(t)] \Delta t$$

The following property models how sensor states are updated. It applies to stimulus s , effect b , and context c (indicated by variable W).

LP7 Generating a sensor state for a world state

If world state W has level X_1

and the sensor state for W has level X_2

then after Δt the sensor state for W will have level $X_2 + \eta [c(\omega_{13}X_1) - X_2] \Delta t$.

$$\text{dss}_W(t)/dt = \eta [c(\omega_{13}\text{ws}_W(t)) - \text{ss}_W(t)]$$

$$\text{ss}_W(t + \Delta t) = \text{ss}_W(t) + \eta [c(\omega_{13}\text{ws}_W(t)) - \text{ss}_W(t)] \Delta t$$

A retrospective ownership state takes into account the prior ownership, the execution of the action, the context, and the sensory representation of the action's effect [see (4) in Sect. 8.2]:

LP8 Generating a retrospective ownership state

If the sensory representation of context c has level X_1 ,

and the sensory representation of effect state e has level X_2

and prior ownership of a for e , c , and s has level X_3

and the execution state for action a has level X_4

and retrospective ownership of a for e , c , and s has level X_5

then after Δt retrospective ownership of a for e , c , and s will have

level $X_5 + \eta [c(\omega_{14}X_1, \omega_{15}X_2, \omega_{16}X_3, \omega_{17}X_4) - X_5] \Delta t$.

$$\text{dros}_{a,e,c,s}(t)/dt = \eta [c(\omega_{14}\text{srs}_c(t), \omega_{15}\text{srs}_e(t), \omega_{16}\text{pos}_{a,e,c,s}(t), \omega_{17}\text{es}_a(t)) - \text{ros}_{a,e,c,s}(t)]$$

$$\text{ros}_{a,e,c,s}(t + \Delta t) = \text{ros}_{a,e,c,s}(t)$$

$$+ \eta [c(\omega_{14}\text{srs}_c(t), \omega_{15}\text{srs}_e(t), \omega_{16}\text{pos}_{a,e,c,s}(t), \omega_{17}\text{es}_a(t)) - \text{ros}_{a,e,c,s}(t)] \Delta t$$

Note that LP8 applies for context self as context, but also to an observed other person B. For an observed other person as context the connection strength ω_{17} in LP8 is assumed 0 or negative; in the simulated scenarios discussed in Sect. 8.4 it was set $\omega_{17} = -1$. The communication to attribute authorship (to any context c) depends on the retrospective ownership state as specified in LP9 [see (6) in Sect. 8.2].

LP9 Communication of ownership awareness

If retrospective ownership of a for e , c , and s has level X_1 ,

and communication of a for e , c , and s has level X_2

then after duration Δt communication of a for e , c , and s will have level $X_2 + \eta [c(\omega_{18}X_1) - X_2] \Delta t$.

$$\text{desc}_{a,e,c,s}(t)/dt = \eta [c(\omega_{18}\text{ros}_{a,e,c,s}(t)) - \text{esc}_{a,e,c,s}(t)]$$

$$\text{esc}_{a,e,c,s}(t + \Delta t) = \text{esc}_{a,e,c,s}(t) + \eta [c(\omega_{18}\text{ros}_{a,e,c,s}(t)) - \text{esc}_{a,e,c,s}(t)] \Delta t$$

LP1 Sensory representation for a sensor state

$$\text{dsrs}_W(t)/dt = \eta [c(\omega_{1\text{ss}_W}(t)) - \text{srs}_W(t)]$$

$$\text{srs}_W(t + \Delta t) = \text{srs}_W(t) + \eta [c(\omega_{1\text{ss}_W}(t)) - \text{srs}_W(t)] \Delta t$$

LP2 Sensory representation for an effect state

$$\text{dsrs}_e(t)/dt = \eta [c(\omega_{2\text{ps}_a}(t), \omega_{20\text{pos}_{a,e,\text{self},s}}(t), \omega_{3\text{ss}_e}(t)) - \text{srs}_e(t)]$$

$$\text{srs}_e(t + \Delta t) = \text{srs}_e(t) + \eta [c(\omega_{2\text{ps}_a}(t), \omega_{20\text{pos}_{a,e,\text{self},s}}(t), \omega_{3\text{ss}_e}(t)) - \text{srs}_e(t)] \Delta t$$

LP3 Preparing for an action

$$\text{dps}_a(t)/dt = \eta [c(\omega_{4\text{srs}_s}(t), \omega_{5\text{srs}_e}(t)) - \text{ps}_a(t)]$$

$$\text{ps}_a(t + \Delta t) = \text{ps}_a(t) + \eta [c(\omega_{4\text{srs}_s}(t), \omega_{5\text{srs}_e}(t)) - \text{ps}_a(t)] \Delta t$$

LP4 Generating a prior ownership state

$$\text{dpos}_{a,e,c,s}(t)/dt = \eta [c(\omega_{6\text{srs}_c}(t), \omega_{7\text{srs}_s}(t), \omega_{8\text{srs}_e}(t), \omega_{9\text{ps}_a}(t)) - \text{pos}_{a,e,c,s}(t)]$$

$$\begin{aligned} \text{pos}_{a,e,c,s}(t + \Delta t) = & \text{pos}_{a,e,c,s}(t) \\ & + \eta [c(\omega_{6\text{srs}_c}(t), \omega_{7\text{srs}_s}(t), \omega_{8\text{srs}_e}(t), \omega_{9\text{ps}_a}(t)) - \text{pos}_{a,e,c,s}(t)] \Delta t \end{aligned}$$

LP5 Action execution

$$\text{des}_a(t)/dt = \eta [c(\omega_{10\text{pos}_{a,e,\text{self},s}}(t), \omega_{11\text{ps}_a}(t)) - \text{es}_a(t)]$$

$$\text{es}_a(t + \Delta t) = \text{es}_a(t) + \eta [c(\omega_{10\text{pos}_{a,e,\text{self},s}}(t), \omega_{11\text{ps}_a}(t)) - \text{es}_a(t)] \Delta t$$

LP6 From action execution to effect state

$$\text{dws}_b(t)/dt = \eta [c(\omega_{12\text{es}_a}(t)) - \text{ws}_e(t)]$$

$$\text{ws}_b(t + \Delta t) = \text{ws}_b(t) + \eta [c(\omega_{12\text{es}_a}(t)) - \text{ws}_e(t)] \Delta t$$

LP7 Generating a sensor state for a world state

$$\begin{aligned} dss_W(t)/dt &= \eta[c(\omega_{13}ws_W(t)) - ss_W(t)] \\ ss_W(t + \Delta t) &= ss_W(t) + \eta[c(\omega_{13}ws_W(t)) - ss_W(t)]\Delta t \end{aligned}$$

LP8 Generating a retrospective ownership state

$$\begin{aligned} dros_{a,e,c,s}(t)/dt &= \eta[c(\omega_{14}srs_c(t), \omega_{15}srs_e(t), \omega_{16}pos_{a,e,c,s}(t), \omega_{17}es_a(t)) - ros_{a,e,c,s}(t)] \\ ros_{a,e,c,s}(t + \Delta t) &= ros_{a,e,c,s}(t) \\ &\quad + \eta[c(\omega_{14}srs_c(t), \omega_{15}srs_e(t), \omega_{16}pos_{a,e,c,s}(t), \omega_{17}es_a(t)) - ros_{a,e,c,s}(t)]\Delta t \end{aligned}$$

LP9 Communication of ownership awareness

$$\begin{aligned} desc_{a,e,c,s}(t)/dt &= \eta[c(\omega_{18}ros_{a,e,c,s}(t)) - esc_{a,e,c,s}(t)] \\ esc_{a,e,c,s}(t + \Delta t) &= esco_{a,e,c,s}(t) + \eta[c(\omega_{18}ros_{a,e,c,s}(t)) - esc_{a,e,c,s}(t)]\Delta t \end{aligned}$$

Box 8.1 Numerical representation of the temporal-causal network model in differential equation format

8.4 Simulation of Example Scenarios

In this section simulations are discussed for a number of example scenarios, which all involve the occurrence of a preparation state for an action a , triggered by some stimulus s . These scenarios relate to phenomena in the literature, as discussed in Sect. 8.2. They have been generated based on the specification in differential equation format shown in Box 8.1. First a scenario is addressed where the prepared action has *satisfactory predicted effects* and therefore is executed; in this case both prior and retrospective self-ownership states occur. Next, a case is considered where the prepared action lacks positive predicted effects, and is therefore not executed: a *no-go decision*, or *vetoing*. Only a rather low prior self-ownership state is developed and no retrospective self-ownership state. In the third case, a *poor action prediction capability* is modeled, which leads to a not very high prior self-ownership state, but sufficient to actually execute the prepared action. In this case no retrospective self-ownership state occurs, as the sensory representation of the effect stays low. In the fourth case, the stimulus triggering the action preparation is the observation of another person performing the action. In this case a low prior *self*-ownership state is generated, but high prior and retrospective *other*-ownership states. This models *mirroring* of and attribution to the other person. Note that the parameter values for the connection strengths used for these different scenarios are not unique. Similar patterns are obtained when they are in a certain range.

8.4.1 Normal Execution and Attribution of an Action

The first case considered describes a situation where the context c is the person itself, and a stimulus s occurs. The action effect e is considered positive for the person. To make things more specific, imagine the following:

- context c is that you are working on your computer
- stimulus s is that you need a certain application P
- action a is clicking on a specific icon I on the desktop of your computer
- effect e of action a is that application P is opened.

The first scenario is as follows:

Scenario 1

- external stimulus s occurs and triggers preparation of action a (to click on icon I)
- based on the preparation state for a the sensory representation of predicted effect e (that application P is opened) of a is generated
- based on this positive predicted effect and the other states a prior self-ownership state for action a (to click on icon I) is generated
- this prior self-ownership state for action a leads to actual execution of action a (mouseclick on I)
- the execution of a affects e in a positive manner (application P opens) and, via sensing, also the sensory representation of e (that application P is opened)
- at the same time the sensory representation of e is suppressed due to the prior self-ownership state
- based on the generated states, after the execution of action a , the person develops a retrospective self-ownership state (for opening application P by clicking on icon I)
- finally the person communicates this self-ownership ('I opened application P by clicking on icon I ').

The simulation of this scenario is shown in Fig. 8.3. Parameter values used (in all scenarios set by hand) can be found in Table 8.3. The step size chosen is

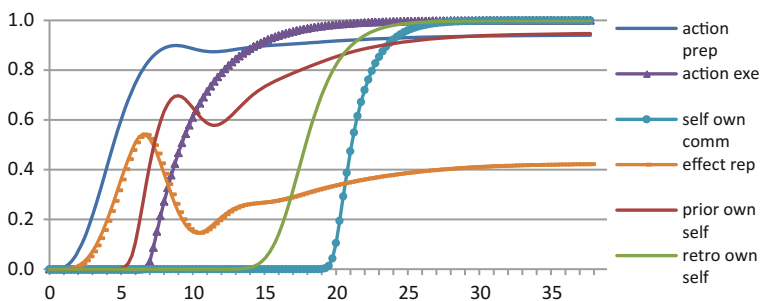


Fig. 8.3 Executing an action with ownership states (scenario 1)

Table 8.3 Parameter values for the first scenario

Connections		Threshold and steepness for state	τ	σ
ω_3	0.5	Action preparation	0.4	4
ω_{2e}	-1	Effect representation	0.2	4
ω_2	0.8	Action execution	1.2	20
ω_4	0.8	Prior self-ownership	3	8
ω_5	0.8	Retrospective self-ownership	3	20
η	0.6/0.3	Self-ownership communication	0.8	40

$\Delta t = 0.25$. All relevant connection strengths not mentioned in this table were chosen 1. The slow value 0.3 for η was applied for external processes (action execution, effect generation and effect sensing) modeled by LP5, LP6, and LP7, and the fast value 0.6 for η for the internal processes modeled by the other LP's.

In Fig. 8.3 it is shown that (after sensing the stimulus), the preparation for action a starts around time point 2, and the representation of the predicted effect e around time point 3. As a result of this, around time point 6 the prior self-ownership state starts to develop, which leads to the execution of the action, starting around time point 7. In the meantime the representation of the action effect e is suppressed (e.g., Blakemore et al. 1999, 2000a, b; Fournier et al. 2002), causing a dip in the graph around time point 10. When the execution of the action a is taking place, the sensing of its effect e in the world has a positive impact on the representation of e from time point 10 on, and the retrospective self-ownership state is developed, starting from around time 15. After this, the communication of the self-ownership takes place from time point 20.

Note that in this case both the prior and the retrospective self-ownership state reach levels close to 1 (prior self-ownership approaching 0.95, and retrospective self-ownership approaching 1). Moreover, note that when the stimulus is taken away, all activation levels will go down to 0, and will come up again when the stimulus reoccurs.

8.4.2 *Vetoing a Prepared Action Due to Unsatisfactory Predicted Effect*

The second case considered describes a situation similar to the previous one (the context c is the person itself, and a stimulus s occurs), but where the action a triggered by stimulus s has an effect e' which is not particularly positive for the person; here a hardly has an impact on effect e which would have been positive. Prediction capabilities are assumed correct in this case, so no high level of e is

correctly predicted for a . For the imagined application opening context this means that it is predicted that clicking on icon I does not open application P but a different application P'. Nevertheless it is assumed that the stimulus triggers preparation for clicking on icon I (maybe because in the past on the same position the icon for program P was placed). For this situation the following variation on the previous scenario is considered:

Scenario 2

- external stimulus s occurs and triggers preparation of action a (to click on icon I)
- based on the preparation state of a only a low level for the sensory representation of predicted effect e of a is generated (no prediction that application P will be opened but that different application P' will be opened)
- based on this low predicted effect e and the other states a low level of a prior self-ownership state for action a is generated (low prior ownership to click on icon I)
- the low prior self-ownership state for a does not lead to actual execution of action a ; the action a can be considered vetoed (no click on icon I)
- the person develops no retrospective self-ownership state for a (no retrospective self-ownership for clicking on icon I)
- the person does not communicate self-ownership for a .

The simulation of this scenario is shown in Fig. 8.4. This scenario was modeled by making the connection strength for the prediction of effect e for action a low: $\omega_2 = 0.3$ instead of 0.8. Values for the other parameters were the same as in Table 8.3.

In Fig. 8.4 it is shown that (after sensing the stimulus), again the preparation for action a starts around time point 2, and the representation of the predicted effect e around time point 3. However, the predicted effect is much lower compared to the previous scenario. As a result of this low prediction, the prior self-ownership state starting to develop around time point 6, also stays at a low level. Therefore the execution of the action also stays very low. Due to these circumstances, no retrospective self-ownership state and no communication of self-ownership occur.

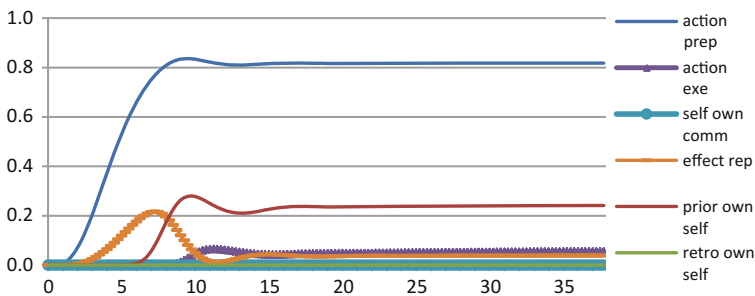


Fig. 8.4 Vetoing an action with no positive prediction (scenario 2)

8.4.3 Effects of Poor Prediction; Schizophrenia Case

The third case considered describes a situation where again the context c is the person itself, and stimulus s occurs. The action effect for action a is e , which in principle is positive for the person, like in the first situation above. However, due to poor prediction capabilities this effect is not (fully) internally predicted. This is what is assumed to happen in patients with schizophrenia, as discussed, for example, in Synofzik et al. (2010) and Voss et al. (2010). For this situation the following scenario is considered:

Scenario 3

- stimulus s occurs and triggers preparation of action a (to click on icon I)
- based on the preparation state for a only a relatively low level of the sensory representation of the predicted effect e of a is generated, due to poor prediction capabilities (low prediction that application P will be opened)
- based on this relatively low predicted effect and the other states a relatively low level of a prior self-ownership state for action a is generated (low prior self-ownership for clicking on icon I)
- this prior self-ownership state level for action a is still sufficient to lead to actual execution of action a (mouseclick on icon I)
- the execution of a affects e (application P is opened) in a positive manner and (via sensing) the sensory representation of e
- the sensory representation of e (that application P is opened) is suppressed to a certain extent due to the (relatively low) prior self-ownership state
- due to the relatively low level for the sensory representation of effect e (and prior self-ownership state) the person develops no retrospective self-ownership state for action a (no retrospective self-ownership for clicking on icon I)
- the person does not communicate self-ownership for action a .

The simulation of this scenario is shown in Fig. 8.5. This scenario was modeled by taking the connection strength for the prediction of effect e for action a moderately low: $\omega_2 = 0.4$. Values for the other parameters were again the same as in

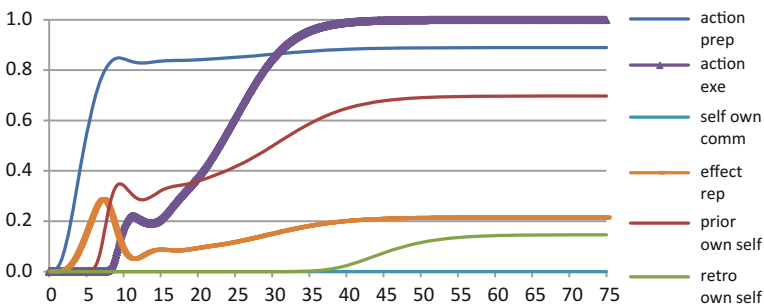


Fig. 8.5 Poor prediction implies no retrospective self-ownership (scenario 3)

Table 8.3. For this case $\Delta t = 0.1$ was chosen instead of 0.25, and the simulation was shown up to time point 75.

In Fig. 8.5 it is shown that as in the previous scenarios, the preparation for action a starts around time point 2, and the representation of the predicted effect e around time point 3. The predicted effect is substantially lower compared to the first scenario, but higher than in the second scenario. As a result of this moderately low prediction, the prior self-ownership state, starting to develop around time point 6, also stays at a moderate level (first around 0.4, later going up to almost 0.7); this is substantially higher than in the second scenario where the lower level led to a veto for the action. Therefore, in contrast to the previous scenario, this level turns out high enough for the execution of the action starting around time point 9. Nevertheless, only a low level of the retrospective self-ownership state is developed (becoming approximately 0.15), and no communication of self-ownership takes place.

8.4.4 *Satisfactory Predicted Effects but Unsatisfactory Actual Effects*

The fourth case considered describes a situation where again the context c is the person itself, and stimulus s occurs. The predicted action effect for action a is e , which in principle is positive for the person, like in the first situation above. However, after executing the action, it turns out that e is not an actual effect of the action. This is another way of modeling a mismatch between prediction and actual outcome. For this situation the following scenario is considered:

Scenario 4

- external stimulus s occurs and triggers preparation of action a (to click on icon I)
- based on the preparation state for a the sensory representation of predicted effect e of a is generated (that application P will be opened)
- based on this positive predicted effect and the other states a prior self-ownership state for action a is generated (to click on icon I)
- this prior self-ownership state for action a leads to actual execution of action a (mouseclick on icon I)
- the execution of a does not affect e in a positive manner (application P is not opening, but a different application P') and; therefore e is not sensed, and no contribution occurs to the sensory representation of e
- at the same time the sensory representation of e is suppressed due to the prior self-ownership state
- based on the generated states, the person develops no retrospective self-ownership state (no retrospective self-ownership for clicking on icon I)
- the person does not communicate self-ownership for action a .

The simulation of this scenario is shown in Fig. 8.6. This scenario was modeled by taking the connection strength for the actual effect e for action a very low:

$\omega_{12} = 0$. Values for the other parameters were again the same as in Table 8.3. For this case $\Delta t = 0.25$ was used.

In Fig. 8.6 it is shown that as in the previous scenarios, the preparation for action a starts around time point 2, and the representation of the predicted effect e around time point 3. The predicted effect is comparable to the first scenario. As a result of this prediction, the prior self-ownership state, develops from around time point 6 on (going up to 0.7). Therefore, this level is high enough for the execution of the action starting around time point 9. However, the expected effect e of the action does not actually occur. Moreover due to the prior ownership state the sensory representation of the effect is suppressed. Therefore immediately after time point 6, the sensory representation of the effect goes down. As a consequence no retrospective self-ownership state is developed, and no communication of self-ownership takes place.

8.4.5 *Mirroring Another Person*

In contrast to the first four scenarios, the fifth case describes a situation where the context c is *another person*, and the stimulus s is the observation of the other person performing action a (person B is clicking on icon I). The action effect for action a is e (application P is opened) and is predicted in a correct manner, as in the first scenario. The scenario for this fourth case is as follows:

Scenario 5

- external stimulus s which is an observed action a performed by another person (by mirroring) triggers preparing action a (to click on icon I)
- based on the preparation state for a the sensory representation of the predicted effect e of a is generated (that application P will be opened)
- based on this predicted effect and the other states (among which the other person as context) a high level of a prior *other*-ownership state for action a is generated, and a low level of a prior *self*-ownership state (other-ownership for clicking on icon I; no self-ownership for clicking on icon I)

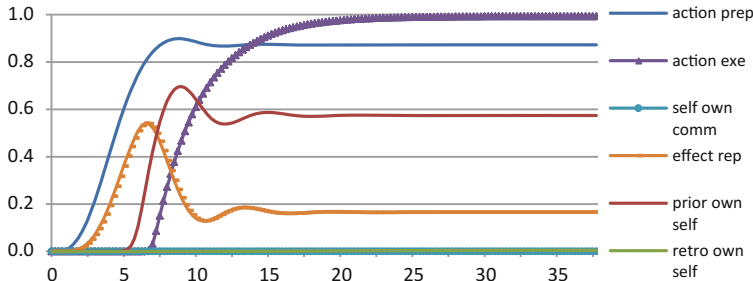


Fig. 8.6 Deviation of actual effect from predicted effect implies no retrospective self-ownership (scenario 4)

- the low prior self-ownership state for action a leads to no actual execution of action a (vetoing; no own mouseclick on icon I)
- as the prior self-ownership state has a low level, not much suppression of the representation of effect e takes place
- based on the generated states, the person develops a retrospective other-ownership state, and no retrospective self-ownership state (retrospective other-ownership for person B opening application P by clicking on icon I)
- finally the person communicates this other-ownership a ('you opened application P by clicking on icon I').

The simulation of this scenario is shown in Fig. 8.7. This scenario was modeled by taking the connection strength ω_4 for mirroring from the specific stimulus representation (observed action) to preparation state 0.5. The connection strength ω_2 for the prediction of effect e for action a is 0.8, as in the first scenario. The threshold and steepness values for prior and retrospective other-ownership states were set 3 and 8, resp. 2.4 and 20. For this case $\Delta t = 0.25$ was used. Values for the other parameters were the same as for the first scenario.

In Fig. 8.7 it is shown that after sensing the observed action, as in the first scenario the preparation for action a starts around time point 2, and the representation of the predicted effect e around time point 3. As a result of this, around time point 6 the prior *other*-ownership state starts to develop, whereas the prior *self*-ownership state stays very low. Therefore execution of a is suppressed. After time point 9 also the retrospective other-ownership state is generated, which leads to communication of other-ownership after time point 12. All other states stay low.

8.5 Relations to Neurological Findings

In this section the question is addressed how the states and dynamical relations of the temporal-causal network model described in Sect. 8.3 relate to neurological states and mechanisms. In particular, this concerns the internal states: sensory representations, preparations, and the prior and retrospective ownership states (in

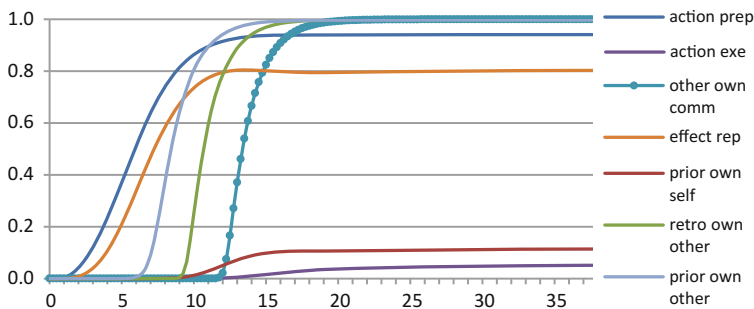


Fig. 8.7 Mirroring another person (scenario 5)

Fig. 8.2 the states in the box). First of all, for the sensory representations of stimulus and context and action effect, these are considered to be related to sensory neurological states. Secondly, the preparation state used is assumed to have a mirroring function as well, so this state may be related to (a group of) mirror neurons (in the classical sense); (e.g., Iacoboni 2008a).

The remaining two types of states, the prior and retrospective ownership states are the most crucial states in the presented model. A prior ownership state has control over the execution of a (prepared) action as one of its most important functions. This type of state has roughly the following behaviour in relation to execution:

- it is active if a prepared action is (to be) executed
- it is not active when no execution (is to) take(s) place.

Interestingly, in recent human single cell recording experiments, specific neurons have been found with activation patterns that have some correlation to execution of an action, in particular, in work reported in Mukamel et al. (2010) and Fried et al. (2011); see also Keysers and Gazzola (2010), Iacoboni (2008a, b) and Iacoboni and Dapretto (2006). For example, Iacoboni (2008b) describes these experiments in patients with epilepsy undergoing pre-surgical evaluation of the foci of epilepsy as follows; see also Iacoboni (2008a, pp. 201–203).

From a total of 14 patients, we have recorded the activity of approximately 500 neurons located in three sectors of the mesial frontal cortex: the ventral and dorsal sectors of the anterior cingulate cortex and the pre-supplementary motor cortex (SMA)/SMA proper complex (Iacoboni 2008b, p. 30).

Some of the main findings are that neurons with mirror neuron properties were found in all sites in the mesial frontal cortex where recording took place (approximately 12 % of all recorded neurons); half of them related to hand-grasping, and the other half to emotional face expressions. For the relation of their activity to execution the following was found:

One-third of mirror neurons had excitatory responses during both action execution and action observation. This is the most typical pattern of firing-rate changes observed in monkeys. One-third of mirror neurons, however, had inhibitory responses during both action execution and action observation. This pattern has also been occasionally observed in monkeys, but much less frequently. The remaining third of mirror neurons in the human frontal cortex had a pattern of firing-rate changes that has never been observed in monkeys, at least not so far (Iacoboni 2008b, p. 30).

It is in this group of mirror neurons that neurons can be found that show behaviour that is similar to the behaviour of the prior ownership state in the presented model as indicated above. A substantial subset of these latter set of neurons...:

...have excitatory responses during action execution and inhibitory responses during action observation. Few of these neurons have the opposite pattern, with decreased firing rate during execution and increased firing rate during observation (Iacoboni 2008b, p. 30).

In Iacoboni (2008a, b) and Iacoboni and Dapretto (2006) such types of neurons are termed *super mirror neurons*, to indicate the control function they may have with respect to the execution of an action. Given the similarity in behaviour, it is this type of neurons that can be considered a suitable candidate as a neurological counterpart of prior ownership states. One aspect to be addressed is the timing aspect: are these super mirror neurons active before the execution, or during or after? In the latter case, they could be considered more suitable as candidates for retrospective ownership states. To obtain more evidence, it would be interesting to find out more about temporal dependencies between the respective neural and execution states, for example, using methods as described in Schippers and Keysers (2011).

For the remaining type of states to be considered, the retrospective ownership states, a more complex picture occurs. Retrospective ownership states have a strong relation to self-monitoring, as also emphasized in Arbib and Mundhenk (2005), Arbib (2007) and Moore and Taggard (2008). For example, in David et al. (2008), a number of cortical and subcortical elements are indicated that may relate as neural correlates to retrospective ownership states:

...several brain areas have been implicated in the sense of agency (...). These include brain regions known to be involved in the motor system such as the ventral premotor cortex (vPMC), the supplementary motor area (SMA and pre-SMA) and the cerebellum as well as regions such as the dorsolateral prefrontal cortex (DLPFC), the posterior parietal cortex (PPC), the posterior segment of the superior temporal sulcus (pSTS) and the insula (...). Unfortunately, the current literature does not yet provide a consistent or clear picture with respect to the exact functions and contributions of these brain regions to the sense of agency (David et al. 2008, p. 529).

In David et al. (2008) a distinction is made into what are called executive functions and supervisory functions, and empirical results in the literature are discussed accordingly.

In a classificatory attempt, the first group of brain regions (e.g., vPMC, SMA, cerebellum) constitutes a network of sensorimotor transformations and motor control, whereas the second group of brain regions rather represents a set of heteromodal association cortices implicated in various cognitive functions. (...) Accordingly, motor system-related regions may subserve 'executive' functions whereas heteromodal associative regions subserve 'supervisory' functions. However, the proposed classificatory, functional distinction remains speculative requiring further empirical validations. In the following, we summarize the available evidence on the neuroscience of agency in an attempt to putting the empirical results into perspective. The presence of different neural correlates might reflect different agency indicators, sub-processes or levels of agency processing (David et al. 2008, pp. 529–530).

So, the possibility is left open that in the end different agency indicators may be distinguished. One of such distinctions may be between unconscious and conscious ownership states; see, for example, Jeannerod (2009) and Schutz-Bosbach et al. (2009).

8.6 Discussion

The temporal-causal network model presented in this chapter incorporates mechanisms for prior and retrospective ownership states, based on principles from recent neurological theories, in particular from Moore and Haggard (2008), Synofzik et al. (2010) and Voss et al. (2010). The contents of this chapter are based on Treur (2011). In the model a prior ownership state is affected by prediction of the effects of the action. Actual execution of the action and sensing of its effects can lead to a retrospective ownership state, in particular, when the sensed effects co-occur with the predicted effects. As a prior ownership state may lead to actual execution of the action, it plays an important role as control of the execution of prepared actions. A retrospective ownership state is the basis for acknowledging authorship of an action, for example, in social context, or in a self-reflection context. Elements of this ownership model have been used in as an important ingredient in the temporal-causal network model for joint decision making described in Chap. 10.

In simulated scenarios it was shown how a number of known phenomena can occur. For example, scenarios were shown for vetoing a prepared action due to unsatisfactory predicted effects, and for mirroring an observed action performed by another person, without imitating the action. Moreover, it was shown how poor action effect prediction capabilities can lead to reduced retrospective ownership states (as, for example, is shown in persons suffering from schizophrenia), and may easily lead to attribution of the self-generated action to another real or imaginary person.

The ownership states are internal cognitive states that are not representations of currently present external things in the way that, for example, sensory representation or belief states are. If the question is posed whether they do represent anything, and if so, what, maybe they can be interpreted as temporal second-order representations or representations of the person's own internal processes, or representations of past or future behaviour, or a combination of these. This representation question is not the focus of this chapter. It may be an interesting topic for further work, for example, in the line of Bosse et al. (2009).

The temporal-causal network model distinguishes itself from existing approaches such as in Wolpert (1997), Frith (1992) and Frith et al. (2000), among others in that (1) instead of comparison of predicted and sensed effects, the predicted and sensed effects are integrated and provide a kind of combined level, as also indicated in, for example Moore and Haggard (2008), Synofzik et al. (2010) and Voss et al. (2010), (2) following Moore and Haggard (2008) a distinction was made between prior and retrospective ownership states, and (3) both self-ownership and other-ownership are covered. These are also differences with approaches, such as for example Hindriks et al. (2011), which do not take the neurological angle as a point of departure, as in the current chapter.

The obtained computational model can be used as a basis for simulation-based training or in gaming or virtual stories. For the first type of application the idea is to develop a virtual patient based on the model so that, for example, a psychiatrist or

psycho-therapist (e.g., during his or her education) can gain insight in the processes in certain types of patients, or it can be used by a therapist to analyse how a certain form of therapy can have its effect on these processes. For the second type of application the idea is to design a system for person-based virtual stories in which, for example, persons with deviations in ownership states play a role (e.g., persons suffering from schizophrenia, and due to that attribute their own actions to other real or imaginary persons), which can be based on the presented model.

Note that the current chapter only addresses the occurrence of ownership states. It does address how weak self-ownership can occur but does not address how in some cases imaginary persons can be created to whom actions are attributed; see for example Collerton et al. (2005) and Samsonovich (2005). In Treur and Umair (2011) it is shown how so-called inverse mirroring enables a person to attribute an action to an imaginary person.

The temporal-causal network model for ownership described in this chapter has been extended to temporal-causal network models also incorporating different awareness states, attention states, a combination of top-down and bottom-up processes, and intentional inhibition of actions in Thilakarathne (2015) and Thilakarathne and Treur (2014, 2015a, b).

References

- M.A. Arbib, Other faces in the mirror: a perspective on schizophrenia. *World Psychiatry* **6**, 75–78 (2007)
- M.A. Arbib, T.N. Mundhenk, Schizophrenia and the mirror system: an essay. *Neuropsychologia* **43**, 268–280 (2005)
- J. Bickle, *Psychoneural Reduction: The New Wave* (MIT Press, Cambridge, Mass, 1998)
- S.-J. Blakemore, C.D. Frith, Self-awareness and action. *Curr. Opin. Neurobiol.* **13**, 219–224 (2003)
- S.-J. Blakemore, C.D. Frith, D.M. Wolpert, Spatio-temporal prediction modulates the perception of self-produced stimuli. *J. Cognit. Neurosci.* **11**, 551–559 (1999)
- S.-J. Blakemore, J. Smith, R. Steel et al., The perception of self-produced sensory stimuli in patients with auditory hallucinations and passivity experiences: evidence for a breakdown in self-monitoring. *Psychol. Med.* **30**, 1131–1139 (2000a)
- S.-J. Blakemore, D.M. Wolpert, C.D. Frith, Why can't you tickle yourself? *NeuroReport* **11**, 11–16 (2000b)
- S.-J. Blakemore, D.M. Wolpert, C.D. Frith, Abnormalities in the awareness of action. *Trends Cognit. Sci.* **6**, 237–242 (2002)
- T. Bosse, C.M. Jonker, J. Treur, Representation for reciprocal agent-environment interaction. *Cognit. Syst. Res. J.* **10**, 366–376 (2009)
- G. Claxton, Why can't we tickle ourselves? *Percept. Mot. Skills* **41**, 335–338 (1975)
- D. Collerton, E. Perry, I. McKeith, Why people see things that are not there: a novel perception and attention deficit model for recurrent complex visual hallucinations. *Behav. Brain Sci.* **28** (6), 737–757 (2005)
- D.F. Collins, T. Cameron, D.M. Gillard, A. Prochazka, Muscular sense is attenuated when humans move. *J. Physiol.* **508**, 635–643 (1998)
- N. David, A. Newen, K. Vogeley, The “sense of agency” and its underlying cognitive and neural mechanisms. *Conscious. Cogn.* **17**, 523–534 (2008)

- I. Feinberg, Efference copy and corollary discharge: Implications for thinking and its disorders. *Schizophr. Bull.* **4**, 636–640 (1978)
- P. Fourneret, F. de Vignemont, N. Franck, A. Slachevsky, B. Dubois, M. Jeannerod, Perception of self-generated action in schizophrenia. *Cogn. Neuropsychiatry* **7**, 139–156 (2002)
- I. Fried, R. Mukamel, G. Kreiman, Internally generated preactivation of single neurons in human medial frontal cortex predicts volition. *Neuron* **69**, 548–562 (2011)
- C.D. Frith, *The Cognitive Neuro-psychology of Schizophrenia* (Lawrence Erlbaum, Hove, UK, 1992)
- C.D. Frith, S. Blakemore, D. Wolpert, Explaining the symptoms of schizophrenia: abnormalities in the awareness of action. *Brain Res. Rev.* **31**, 357–363 (2000)
- P. Haggard, Human volition: towards a neuroscience of will. *Nat. Neurosci. Rev.* **8**, 934–946 (2008)
- R.D. Hawkins, E.R. Kandel, Is there a cell-biological alphabet for simple forms of learning? *Psychol. Rev.* **91**, 375–391 (1984a)
- R.D. Hawkins, E.R. Kandel, Steps toward a cell-biological alphabet for elementary forms of learning, in *Neurobiology of Learning and Memory*, ed. by G. Lynch, J.L. McGaugh, N.M. Weinberger (Guilford Press, New York, 1984b), pp. 385–404
- K. Hindriks, P. Wiggers, C.M. Jonker, W. Haselager, Towards a computational model of the self-attribution of agency, in *Proceedings of the 24th International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems, IEA/AIE'11, Part I*. Lecture Notes in AI, vol. 6703 (Springer Verlag, 2011), pp. 295–305
- M. Iacoboni, *Mirroring People: The New Science of How We Connect with Others* (Farrar, Straus & Giroux, 2008a)
- M. Iacoboni, Mesial frontal cortex and super mirror neurons. *Behav. Brain Sci.* **31**, 30 (2008b)
- M. Iacoboni, M. Dapretto, The mirror neuron system and the consequences of its dysfunction. *Nat. Rev. Neurosci.* **7**, 942–951 (2006)
- M. Jeannerod, The sense of agency and its disturbances in schizophrenia: a reappraisal. *Exp. Brain Res.* **192**, 527–532 (2009)
- C. Keysers, V. Gazzola, Social neuroscience: mirror neurons recorded in humans. *Curr. Biol.* **20**, 253–254 (2010)
- J. Moore, P. Haggard, Awareness of action: inference and prediction. *Conscious. Cogn.* **17**, 136–144 (2008)
- R. Mukamel, A.D. Ekstrom, J. Kaplan, M. Iacoboni, I. Fried, Single-neuron responses in humans during execution and observation of actions. *Curr. Biol.* **20**, 750–756 (2010)
- Samsonovich, A.V. (2005) Hallucinating objects versus hallucinating subjects. *Behavioral and Brain Sciences* (2005) 28:6, pp. 772–773
- M.B. Schippers, C. Keysers, Mapping the flow of information within the putative mirror neuron system during gesture observation. *Neuroimage* **57**, 37–44 (2011)
- S. Schutz-Bosbach, A. Avenanti, S.M. Aglioti, P. Haggard, Don't do it! cortical inhibition and self-attribution during action observation. *J. Cogn. Neurosci.* **21**, 1215–1227 (2009)
- M. Synofzik, P. Thier, D.T. Leube, P. Schlotterbeck, A. Lindner, Misattributions of agency in schizophrenia are based on imprecise predictions about the sensory consequences of one's actions. *Brain* **133**, 262–271 (2010)
- D.J. Thilakaratne, Modelling of situation awareness with perception, attention, and prior and retrospective awareness. *Biol. Insp. Cognit. Arch. Elsevier* **12**, 77–104 (2015)
- D.J. Thilakaratne, J. Treur, Modelling the dynamics of emotional awareness, in *Proceedings of the 21th European Conference on Artificial Intelligence, ECAI'14*, eds. by T. Schaub, G. Friedrich, B. O'Sullivan, vol. 263, (IOS Press, Frontiers in Artificial Intelligence and Applications, 2014), pp. 885–890
- D.J. Thilakaratne, J. Treur, Computational cognitive modeling of action awareness: prior and retrospective. *Brain Inform.* **2**(2), 77–106 (2015a)
- D.J. Thilakaratne, J. Treur, Modelling intentional inhibition of actions. *Biol Inspir. Cognit. Arch. Elsevier* **14**, 22–39 (2015b)

- J. Treur, A cognitive agent model incorporating prior and retrospective ownership states for actions, in *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence, IJCAI'11*, ed. by T. Walsh (2011), pp. 1743–1749
- J. Treur, M. Umair, A cognitive agent model using inverse mirroring for false attribution of own actions to other agents, in: *Proceedings of the 24th International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems, IEA/AIE'11*. Lecture Notes in Artificial Intelligence, vol. 6704 (Springer Verlag, 2011), pp. 109–119
- M. Voss, J. Moore, M. Hauser, J. Gallinat, A. Heinz, P. Haggard, Altered awareness of action in schizophrenia: a specific deficit in predicting action consequences. *Brain* **133**, 3104–3112 (2010)
- L. Weiskrantz, J. Elliot, C. Darlington, Preliminary observations of tickling oneself. *Nature* **230**, 598–599 (1971)
- D.M. Wolpert, Computational approaches to motor control. *Trends Cognit. Sci.* **1**, 209–216 (1997)

Chapter 9

How Empathic Are You

Displaying, Regulating, and Learning Adaptive Social Responses

Abstract Differences in social responses of individuals can often be related to differences in functioning of certain neurological mechanisms. Based on a Network-Oriented Modeling approach, a temporal-causal network model has been developed that is capable of showing different types of social response patterns according to such mechanisms, adopted from theories on mirror neuron systems, emotion integration, emotion regulation, and empathy. Within the model also adaptive capabilities have been incorporated, showing how learning of social response patterns can take place. The adaptive temporal-causal network model provides a basis for human-like social response patterns of virtual agents in the context of simulation-based training (e.g., for training of physicians or therapists), gaming, or for generation of virtual stories.

9.1 Introduction

Human social interaction often goes beyond verbal exchange of information. For example, to obtain and display forms of mutual empathic understanding, both verbal and nonverbal interaction play a role. Such forms of understanding have been recognized not only to be important to maintain personal relationships, but also in professional relationships, for example, between a teacher and a student, between a counselor and a client, or between a physician and a patient. To monitor such social interactions of professionals, and, when desired, to improve their capabilities in these interactions, specific means and training facilities have been or are being developed. Examples from the medical area are Bonvicini et al. (2009), Hojat (2007, 2009), Suchman et al. (1997), Tulskey et al. (2011), Zimmermann et al. (2011). As, for example, discussed in Tulskey et al. (2011) computer support environments for training purposes may provide a useful contribution to this field. However, to be able to develop environments of good quality, insight in the mechanisms underlying such social interaction is important.

In recent years neurological mechanisms have been discovered that describe how, for example, direct nonverbal contagion of emotions (e.g., responding to a

smile) may take place between persons. Within neuroscience the study of mechanisms behind social interaction has led to a fast developing new discipline called Social Neuroscience (e.g., Cacioppo and Berntson 2005; Cacioppo et al. 2006; Decety and Cacioppo 2010; Harmon-Jones and Winkielman 2007). Examples of processes and mechanisms identified as important for social interaction are mirror neuron systems, self-other distinction, emotion integration, emotion regulation and empathy. Such mechanisms provide a useful point of departure to design biologically plausible computational models that offer a wide human-like social interaction repertoire. Here the concept of mirror neuron is a central concept relating to the other mechanisms as well. Mirror neurons are neurons with both a function of preparing, and of mirroring a similar state of another person; e.g., Iacoboni (2008a), Rizzolatti and Sinigaglia (2008), Pineda (2009).

The collection of mechanisms considered in this paper has resulted from in-depth neurological investigations of deficits in social interaction. In the neurological literature these are put forward as the mechanisms that show impairments for persons with such deficits in social interaction. The contribution of such mechanisms to social functioning is usually studied by comparing a group persons that do not show adequate social interaction with a control group of persons with typical social functioning. Within the natural, human population, substantial differences in social behaviour between different persons occur. Some of the specific types of social interaction are considered to be 'autistic' to a certain extent, and the persons displaying them are sometimes diagnosed as having some form of an Autism Spectrum Disorder (ASD); e.g., Richer and Coates (2001), Frith (2003).

Based on a Network-Oriented Modeling approach a temporal-causal network model has been developed that integrates computational formalisations of mechanisms for mirroring, self-other distinction, emotion integration, emotion regulation, and empathy put forward in the recent neurological literature as crucial for adequate social interaction. These mechanisms have been incorporated in the model in an abstracted form; the network model can be considered a computational model inspired by these neurological mechanisms. Given the use of neurological mechanisms as a point of departure, a biologically plausible model results that can be used as a basis for the development of applications, for example, in the context of simulation-based training, gaming or virtual stories. Such applications can concern software environments using virtual agents based on the model presented here with built-in parameters representing personal characteristics. This does not only allow settings for human-like agents that model an idealised, perfect form of social interaction, but also settings that model different forms of imperfection in social interaction as occurring in the natural human population. In particular, such environments can be helpful in training professionals such as physicians in their social interaction.

It is often claimed that the mirroring mechanism is not (fully) present at birth, but has to be shaped by experiences during lifetime; for example, Catmur et al. (2007), Iacoboni (2008a, b), Keysers and Perrett (2004). For persons (in particular children) with low or no social responses, it is worth while to offer them training sessions in imitation so that the mirror neuron system and the displayed social

responses may improve. This indeed turns out to work, at least for the short term, as has been reported in, for example Field et al. (2001), Ingersoll et al. (2007). Thus evidence is obtained that the mirror neuron system has a certain extent of plasticity due to some learning mechanism. In Keysers and Perrett (2004) it is argued that Hebbian learning (e.g., Gerstner and Kistler 2002; Hebb 1949) is a good candidate for such a learning mechanism.

In this chapter a Hebbian learning mechanism is adopted to obtain an adaptive temporal-causal network model showing plasticity of the person's mirror neuron system. The model realises learning of social behaviour (in particular, empathic social responses), depending on a combination of innate personal characteristics and the person's experiences over time obtained in social context. A person's experiences during lifetime may concern self-generated experiences (the person's responses to other persons encountered) or other-generated experiences (other persons' responses to the person). By varying the combination of innate characteristics and the social context offering experiences, different patterns of learning and unlearning of socially responding to other persons are displayed.

In this chapter, after some background in the relevant neurological mechanisms in the literature is discussed in Sect. 9.2, the design of the temporal-causal network model is presented in Sect. 9.3. In Sect. 9.4 an exploration is presented illustrated by a number of simulation results and (emerging) properties shown by the simulated patterns. In Sect. 9.5 the adaptive model for Hebbian learning of the social behaviour is presented. In Sect. 9.6 some simulation results are discussed, for different characteristics and social contexts. Finally, Sect. 9.7 is a discussion.

9.2 Neurological Background

In this section a review is presented of theories in the social-neurological literature about mechanisms relevant to social interaction. Each subsection describes one of the mechanisms and indicates a different hypothesis about causes of deficits in social interaction due to malfunctioning of that mechanism.

9.2.1 *Mirror Neurons*

It has been found that certain preparation states for actions or for expressing body states (at the neural level related to *mirror neurons*) have multiple functions, not only the function of preparing, but also the function of *mirroring* a similar state of another person; e.g., Iacoboni (2008a), Rizzolatti and Sinigaglia (2008), Pineda (2009), Fried et al. (2011), Keysers and Gazzola (2010), Mukamel et al. (2010). Neurological evidence for specific impairments due to reduced activation of mirror neurons in persons with ASD is reported in, e.g., Dapretto et al. (2006), Iacoboni (2008a), Williams et al. (2001). For example, in Dapretto et al. (2006, p. 30) it is

reported that children with ASD show reduced mirror neuron activity when observing emotional expressions, compared to typically developing children, and the hypothesis is put forward that early dysfunction in the mirror neuron system is at the core of social deficits observed in persons with ASD. This points at the mirror neuron system as a first mechanism which is important for social interaction. Reduced functioning of the mirror neuron system is a first hypothesis about causes of deficits in social interaction.

The functional meaning of activation of mirror neurons (e.g., preparing or mirroring or both) may be strongly context-dependent: in which cases is their activation meant to lead to actual execution of the action, and in which cases it is not. A specific subset of neurons has been found that seem to be able to provide such a context; this is discussed next.

9.2.2 *Control and Self-other Distinction*

Suitable forms of context can be defined at the neurological level based on what sometimes are called *control neurons* or *super mirror neurons* (Iacoboni 2008a, pp. 196–203; 2008b; Brass and Spengler 2009). These are neurons which were suggested to have a function in control (allowing or suppressing) action execution after preparation has taken place. In single cell recording experiments with epileptic patients, cells were found that are active when the person prepares an own action that is executed, but shut down when the action is only observed, which leads to the hypothesis that these cells may be involved in the functional distinction between preparation state generated in order to actually perform the action, and a preparation state generated to interpret an observed action (or both, in case of imitation). More specifically, this has been shown in work reported in Mukamel et al. (2010), Fried et al. (2011); see also Keysers and Gazzola (2010), Iacoboni (2008a, b), Iacoboni and Dapretto (2006). For example, Iacoboni (2008b) describes these experiments in 14 patients with epilepsy undergoing pre-surgical evaluation of the foci of epilepsy; see also Iacoboni (2008a, pp. 201–203). Some of the main findings are that neurons with mirror neuron properties were found in all sites in the mesial frontal cortex where recording took place (approximately 12 % of all recorded neurons); half of them related to hand-grasping, and the other half to emotional face expressions. A subset of neurons was found that show behaviour that relate to execution of the action: they have excitatory responses during action execution and inhibitory responses during action observation (Iacoboni 2008b, p. 30). In Iacoboni (2008a, b), Iacoboni and Dapretto (2006) such types of neurons indicate the control function with respect to the execution of an action. In Iacoboni (2008a, pp. 201–202) it is also described that some of such cells are sensitive to a specific person, so that an observed action can also be attributed to the person that was observed (self-other distinction). In Brass and Spengler (2009) and Hamilton et al. (2007) it is suggested that the types of social interaction seen in persons with ASD can be related to reduced self-other distinction and control of imitation.

9.2.3 *Emotion Integration*

The integration of affective processes in cognitive processes (e.g., Pessoa 2008; Phelps 2006) is another type of mechanism that is assumed to play an important role in social interaction. According to Damasio (1999) sensory representations of stimuli usually induce responses with associated emotions in the form of preparations for modified body states. Activation of such preparation states lead to further mental processing via an *as-if body loop* from preparation state to emotions felt based on sensory representation of body states associated to the prepared action. Conversely, it is assumed that the preparation for the response is also affected by the level of feeling the emotion in the form of the sensory representation of the body state. Thus reciprocal causation relations exist between emotions felt and preparations for actions, which realises *integration of emotion* in preparation of actions; see also Damasio (2003), Bosse et al. (2012). In Grezes and de Gelder (2009), Grezes et al. (2009) the role of emotion integration is emphasized, referring to brain areas such as Superior Temporal Sulcus and Amygdala and their connectivity. In Grezes and de Gelder (2009, pp. 73–74) it is put forward that studies provide evidence that in autistic subjects this reduced connectivity may result in the mirror mechanism (although by itself well functioning) being dissociated from socio-affective capabilities.

9.2.4 *Enhanced Sensory Processing Sensitivity and Emotion Regulation*

A fourth mechanism affecting social interaction is regulation to compensate for enhanced *sensory processing sensitivity*. For example, in Baker et al. (2008, pp. 867–868) it is put forward that dysfunction in processing sensory information results in deviant behaviours to (down)regulate stimulation from the environment; see also Hofsten and Gredebäck (2009). This hypothesis has a long history, going back, for example, to Hutt et al. (1964) and Tinbergen et al. (1972), who compared ASD-related behaviours to stereotyped and avoidance behaviours shown by animals when placed in stressful circumstances. During this long history not all of the several claims made in this direction have been confirmed. Specific difficulties are not only the many different ways and degrees in which ASD-related phenomena occur in different persons, but also the adaptation by internal emotion regulation mechanisms employed to compensate for deviations in sensory processing sensitivity.

In Gross (1998, 2001, 2007), Goldin et al. (2008) a process model of *emotion regulation* is described. Emotion regulation is taken as including all of the conscious and nonconscious strategies used to increase, maintain, or decrease one or more components of an emotional response. The considered emotional responses have *experiential* (subjective feeling of the emotion), *behavioral*, and *physiological*

components (responses such as heart rate and respiration). Regulation strategies are used, differentiated as antecedent-focused strategies and response-focused strategies; see also Chap. 3, Sect. 3.3. *Antecedent-focused strategies* are applied in the process preparing for responses before they are fully activated. *Response-focused strategies* are applied to the actual emotional response, when a response which is already underway is modulated. Gross distinguishes four different types of antecedent-focused emotion regulation strategies: *situation selection*, *situation modification*, *attentional deployment* and *cognitive change*.

Situation selection occurs when a person chooses for a situation that is expected to generate the emotional response level the person wants to have for a certain emotion. For example, a person can go to a party instead of staying home instead, because at the party someone will be met with a positive effect on feeling happy. This is an example of up-regulating one's emotion (happiness). An example of situation selection to down-regulate one's emotion (anger) is avoiding some annoying person. Situation modification means that a person modifies an existing situation so as to obtain a different level of emotion. For instance, when watching a thriller on television, one may zap to another channel when the 'thrill' becomes too strong. Attentional deployment is shifting attention to a certain aspect, for example, closing your eyes when watching an exciting penalty shoot-out. Cognitive change is selecting a specific cognitive meaning to an event. A specific type of cognitive change, which is aimed at down-regulating emotion, is *reappraisal*: the individual reappraises or cognitively re-evaluates a potentially emotion-eliciting situation in terms that decrease its emotional impact (Gross 2001). An example of reappraisal is a case when a person performs bad and blames other circumstances, instead of his own efforts. Response modulation is applied after the emotion response tendencies have been generated: a person tries to affect the response tendencies becoming a behavioral response. A specific type of response modulation, aimed at down-regulating, is *suppression* which means that an individual inhibits ongoing expressive behavior (Gross 2001).

In the specific case of enhanced sensitivity for certain types of stimuli, compensation can take place by forms of emotion regulation by avoiding situations or aspects of situations in which these stimuli occur, or focus attention differently, and/or by suppressing the own bodily response. Such regulation may not only diminish or even eliminate or overcompensate phenomena, which makes them hard to observe in experiments, but as it typically is a cyclic adaptive process it also makes it difficult to attribute causality.

In recent years the perspective of enhanced sensory processing sensitivity has become a quite active area of research; see for example, Baker et al. (2008), Crane et al. (2009), Gepner and Féron (2009), Lane et al. (2010), Smith (2009). Using eye trackers that have become widely available, much work focuses on gaze fixation or gaze aversion behaviour in relation to over-arousal due to enhanced sensitivity for sensory processing of face expressions, in particular in the region of the eyes; e.g., Corden et al. (2008), Kirchner et al. (2011), Kylliäinen and Hietanen (2006), Neumann et al. (2006), Spezio et al. (2007). To get rid of arousal which is experienced as too strong, as a form of antecedent-focused regulation (in particular,

attentional deployment) the gaze can be taken away from the observed face or eyes (gaze aversion). According to this perspective, gaze aversion and showing an expressionless face and (monotonous) voice, as often occur in persons with ASD, can be viewed as forms of regulation of the level of arousal, which otherwise would be experienced as too overwhelming, and disturbing for the other mental processes.

9.2.5 *Empathic Responses*

Developing empathy is an important process as a basis for social interaction. In De Vignemont and Singer (2006), Singer and Leiberg (2009) the following four elements of the process to develop empathy are formulated (see also Chap. 7, Sect. 7.3.2):

- (1) Presence of an affective state in a person
- (2) Isomorphism of the person's own and the other person's affective state
- (3) Elicitation of the person's affective state upon observation or imagination of the other person's affective state
- (4) Knowledge of the person that the other person's affective state is the source of the person's own affective state

The neurological mechanisms to obtain empathy involve mirror neurons, self-other distinction and emotion integration (as described in Sects. 9.2.1–9.2.3 above). Given an affective state in another person (1), mirror neurons (see Sect. 9.2.1) and emotion integration by as-if body loops (Sect. 9.2.3) form a mechanism that generates an own affective state isomorphic with the other person's affective state (2), thereby using observation or imagination of the other person's expressions (3). Moreover, by self-other distinction (see Sect. 9.2.2), knowledge is obtained that the other person is the source of this affective state (4).

These elements and underlying mechanisms can be considered as a basis of developing an internal state of 'having empathy'. However, within social interaction, it is not only important that this occurs as an internal state, but also that this is displayed to the other person. Such an interaction does not only involve displaying the emotion felt ('showing feeling') but also displaying the fact of knowing that it concerns the emotion of the other person ('showing knowing'). Therefore, such a 'displayed empathy' or an 'empathic response', may involve:

- (a) Showing the same emotion as the other person
- (b) Telling that the other person has this emotion

Assuming true, faithful bodily and verbal expression, these two criteria (a) and (b) are entailed by the four criteria of empathy formulated in De Vignemont and Singer (2006), Singer and Leiberg (2009). For example, if it is assumed that the affective state in (1) is shown to the other person by expressing it nonverbally

and/or verbally, then (1) and (2) entail (a). Moreover, if it is assumed that the knowledge in (4) is communicated, then (4) entails (b).

It is generally acknowledged that showing empathy is important in professional relations, for example for physicians; e.g., Bonvicini et al. (2009), Hojat (2007, 2009), Suchman et al. (1997), Tulsy et al. (2011), Zimmermann et al. (2011). The items (a) and (b) will be illustrated for this context. In Suchman et al. (1997, p. 679) the following is one of the example dialogues discussed:

Example 1

PHYSICIAN: How do you feel about the cancer—about the possibility of it coming back?

PATIENT: Well, it bothers me sometimes but I don't dwell on it. But I'm not as cheerful about it as I was when I first had it. I just had very good feelings that everything was going to be all right, you know. But now I dread another operation. [empathic opportunity]

PHYSICIAN: You seem a little upset; you seem a little teary-eyed talking about it. [empathic response]

Note that this is only a partial representation of the social interaction: it is only a linguistic representation of the interaction that does not show the nonverbal expressions of the physician that may have been there accordingly. Such positive example dialogues are contrasted to dialogues where the physician misses the opportunity to show an empathic response, such as the following one (Suchman et al. 1997, p. 679):

Example 2

PHYSICIAN: Does anybody in your family have breast cancer?

PATIENT: No.

PHYSICIAN: No?

PATIENT: Now I just start [unintelligible] after I had my hysterectomy. I was taking estrogen, right?

PHYSICIAN: Yeah?

PATIENT: You know how your breast get real hard and everything? You know how you get sorta scared? [empathic opportunity]

PHYSICIAN: How long were you on the estrogen? [empathic opportunity terminator, missed empathic opportunity]

PATIENT: Oh, maybe about 6 months.

In Example 1 the response at least satisfies (b), and when nonverbal expressions are assumed accordingly it satisfies both (a) and (b). When in Example 1 it would be assumed that the physician keeps a nonexpressive pokerface, it does not satisfy (a). The response in Example 2 does not satisfy (b) and when nonverbal expressions are assumed absent accordingly also not (a). When in Example 2 it would be assumed that the physician still expresses the emotion, it would satisfy (a).

9.3 The Temporal-Causal Network Model

In this section the temporal-causal network model will be described in detail. First an overview will be given, and subsequently the different parts of the model will be addressed: sensory representations, preparations, mirroring and control (sometimes called super mirroring), expressing body states, communication and gaze, maintaining body state and gaze, and generating sensor states.

9.3.1 *Conceptual Representation of the Model*

The theories described in Sect. 9.2 above each point at a different mechanism that is important for social interaction. To obtain adequate social interaction, all of these mechanisms have to function well in conjunction. More specifically, the following theories described in Sect. 9.2 were taken into account in designing the temporal-causal network model:

- mirror neuron systems; e.g., Dapretto et al. (2006), Iacoboni (2008a)
- control neurons with self-other distinction and control function; e.g., Iacoboni (2008a), Brass and Spengler (2009)
- emotion integration; e.g., Grèzes and de Gelder (2009), Grèzes et al. (2009)
- regulation of enhanced sensory processing sensitivity, in particular for face expressions; e.g., Neumann et al. (2006), Spazio et al. (2007), Baker et al. (2008), Corden et al. (2008)
- empathic responding using mirror neurons, self-other distinction and emotion integration; e.g., De Vignemont and Singer (2006), Singer and Leiberg (2009)

A reasonable perspective is that all of the mechanisms as put forward play their role in social interaction in an integrative manner, and if one of them is not functioning well, this may lead to specific deficits in social functioning. Based on this view, in the design of the computational model below an integrative approach has been followed where for each of the mechanisms a computational formalisation was included in the model, and integrated with the computational formalisations of the other mechanisms. When all of the mechanisms work well, this results in adequate social functioning, but when one or more of them do not work well this easily leads to deficits in social interaction. For each of the computational formalisations of the mechanisms such malfunctioning can be specified by specific parameter settings.

So, the elements described above have been exploited in an integrative manner in the presented temporal-causal network model. Thus a human-like model is obtained that, depending on its settings is able to show different types of social response patterns, for example, the type of responses of the physician discussed in Sect. 9.2.5 for the social interaction between a physician and a patient (see Example 1). More specifically, the temporal-causal network model designed incorporates mirroring,

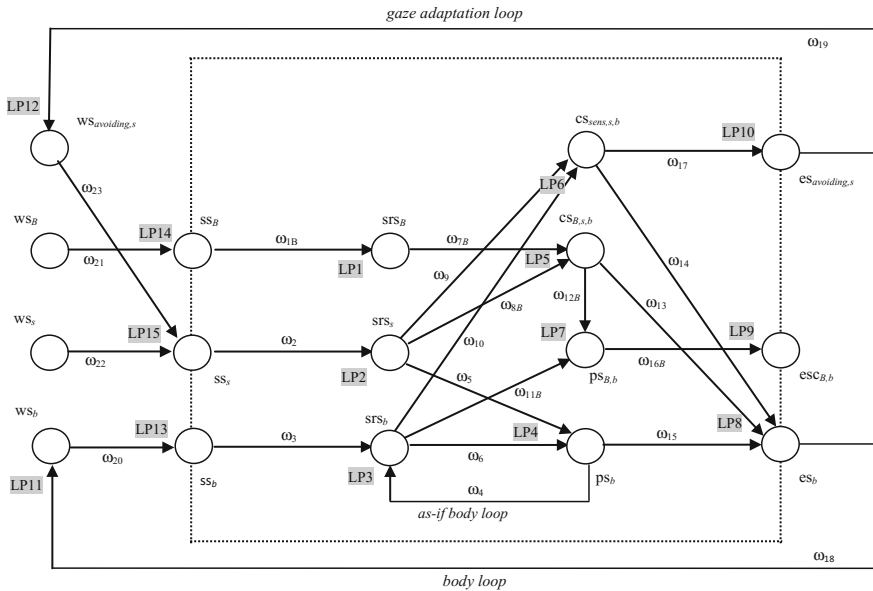


Fig. 9.1 Overview of the temporal-causal network model; see also Table 9.1

self-other distinction and control, emotion integration, gaze adaptation as a form of emotion regulation to compensate for enhanced sensory processing sensitivity, and empathic responding; see Fig. 9.1 for an overview. Here the following notations are used:

- ws world states
- ss sensor states
- srs sensory representation states
- ps preparation for a body state
- cs control state
- es execution states
- esc execution states for (expression of) communication

The connections between the states have weights indicated by ω with subscripts. Furthermore, labels LPi refer to the corresponding detailed dynamic property specification presented below. States that relate to the physical world such as body states, sensor states and effector states are modeled in an abstract form. They may be related to any specific physical mechanisms of choice.

Note that in the causal graph of the model three loops occur: the body loop to adapt the body, the as-if body loop to adapt the internal body map, and the gaze adaptation loop to regulate the enhanced arousal. The effect of these loops is that for any new external situation encountered, in principle, a (numerical) approximation process may take place until the internal states reach an equilibrium (assuming that the situation does not change too fast). However, as will be discussed in Sect. 9.4, it

is also possible that a (static) external situation does not lead to an equilibrium, but to periodic oscillations.

Modeling causal relations discussed in neurological literature in the manner as presented here does not take specific neurons into consideration but uses more abstract cognitive or mental states. In this way abstraction takes place by lifting neurological knowledge to a mental (cognitive/affective) modeling level. For more details of the modeling perspective, see Chaps. 1 and 2.

The temporal-causal network model has been numerically represented in differential equation format. In the model s denotes a stimulus (e.g., a smiling face of another person B , or the tears of the patient in Example 1 in Sect. 9.2.5), b a body state (e.g., a responsive smile or sad face) and B a person (another person or the person *self*). A control state can either refer to a person B , or to enhanced sensory processing sensitivity, indicated by *sens*. Note that, following Damasio (1999), a body state b is used as a label to indicate an emotion, and srs_b the feeling (sensory representation) of the emotion. Communication of b to B means communication that the person *self* knows that B feels b ; e.g., the last line of Example 1 in Sect. 9.2.5:

You seem a little upset; you seem a little teary-eyed talking about it.

The gaze adaptation has been modeled using the notion *avoiding* s for gaze direction in relation to s ; this denotes a specific gaze direction in an area avoiding s .

Connections between states (the arrows in Fig. 9.1) have weights, as indicated in Table 9.1 and in Fig. 9.1. A weight ω_k may depend on a specific stimulus s , and body state b involved, and on a person B (self or another person), when this is indicated by an index B . It usually has a value between 0 and 1, but for suppressing effects it can also be negative. In the column indicated by LP a reference is made to the (temporally) Local Property (LP) that specifies the update dynamics of the activation value of the ‘to state’ based on the activation levels of the ‘from states’; see below.

Table 9.1 Overview of the connections, their weights, and their explanations; see also Fig. 9.1

From states	To state	Weight name	Connection name	LP	Explanation
ss_B	srs_B	ω_{1B}	Representing B	LP1	Representing an person B from sensing B
ss_s	srs_s	ω_2	Representing s	LP2	Representing a stimulus s (e.g., another person B ’s smile or tears)
ss_b ps_b	srs_b	ω_3 ω_4	Representing b Predicting b	LP3	Representing a <i>body map</i> for b : emotion b felt (e.g., own smile) • from sensing own body state b • via <i>as-if body loop</i> from preparation for body state b

(continued)

Table 9.1 (continued)

From states	To state	Weight name	Connection name	LP	Explanation
srs_s srs_b	ps_b	ω_5 ω_6	Responding Amplifying	LP4	Preparing for <i>body state b</i> : emotional response <i>b</i> (e.g., own smile or sad face) • via <i>mirroring</i> from represented stimulus <i>s</i> (e.g., smile of <i>B</i>) • via <i>emotion integration</i> from emotion <i>b</i> felt
srs_B srs_s	$cs_{B,s,b}$	ω_{7B} ω_{8B}	Monitoring <i>B</i> Monitoring <i>s</i> for self-other distinction	LP5	<i>Control state</i> for <i>self-other distinction</i> • from represented person <i>B</i> • from represented stimulus <i>s</i> (e.g., smile or tears of <i>B</i>)
srs_s srs_b	$cs_{sens,s,b}$	ω_9 ω_{10}	Monitoring <i>s</i> for sensitivity Monitoring <i>b</i>	LP6	<i>Control state</i> for <i>enhanced sensitivity</i> • from represented stimulus <i>s</i> (e.g., smile or tears of <i>B</i>) • from emotion <i>b</i> felt
srs_b $cs_{B,s,b}$	$ps_{B,b}$	ω_{11B} ω_{12B}	Communication response Controlling communication	LP7	Preparing <i>communication</i> to <i>B</i> (e.g., ‘you feel <i>b</i> ’) • via <i>emotion integration</i> from emotion <i>b</i> felt • <i>controlled by control state</i> for <i>B</i>
$cs_{self,s,b}$ $cs_{sens,s,b}$ ps_b	es_b	ω_{13} ω_{14} ω_{15}	Controlling response Suppressing response Executing response	LP8	Expressing <i>body state b</i> (e.g., own smile) • <i>controlled by control state</i> for <i>self</i> • <i>controlled by control state</i> for enhanced sensitivity • from preparation state for <i>b</i>
$cs_{B,b}$	$esc_{B,b}$	ω_{16B}	Executing communication	LP9	Expressing <i>communication</i> (e.g., ‘you feel sad’)
$cs_{sens,s,b}$	$es_{avoiding,s}$	ω_{17}	Executing avoidance	LP10	Expressing <i>gaze</i> , <i>controlled by a control state</i> for enhanced sensitivity
es_b	ws_b	ω_{18}	Effectuating <i>b</i>	LP11	Effectuating actual <i>body state</i>
$es_{avoiding,s}$	$ws_{avoiding,s}$	ω_{19}	Effectuating avoidance	LP12	Effectuating actual <i>gaze</i>
ws_b	ss_b	ω_{20}	Sensing <i>b</i>	LP13	Sensing body state <i>b</i>
ws_B	ss_B	ω_{21}	Sensing <i>B</i>	LP14	Sensing a person <i>B</i>
ws_s $ws_{avoiding,s}$	ss_s	ω_{22} ω_{23}	Sensing <i>s</i> Suppressing sensing of <i>s</i>	LP15	Sensing stimulus <i>s</i> • from world state <i>s</i> • <i>regulated by gaze state</i> avoiding <i>s</i>

By varying the connection strengths, different possibilities for the social interaction repertoire offered by the model can be realised. Emotion integration in other states takes place by using a connection from srs_b to these other states: in LP4, LP6, and LP7.

Reduced emotion integration can be expressed by low weights $\omega_6, \omega_{10}, \omega_{11B}$ for these connections. Similarly, low values for ω_5 in LP4, resp. ω_{7B}, ω_{8B} in LP5 can be used to achieve reduced mirroring, resp. control, and higher values for ω_9, ω_{10} in LP6 indicate enhanced sensory processing sensitivity. Below, each of the local dynamic properties is described in more detail as a semiformal description and as a differential equation, summarized in Box 9.1.

9.3.2 Numerical Representation of the Temporal-Causal Network Model

During processing, each state has an activation level represented by a real number between 0 and 1. Parameter η is a speed factor, which determines how fast a state is changing, based on impact received from other states connecting to it, and $c(\dots)$ denotes a combination function used to aggregate different impacts on the same state.

The properties LP1–LP3 describe how sensory representations are generated for a person B , stimulus, and body state.

LP1 Sensory representation of an external person B

If a person B is sensed with level X_1 ,

and the sensory representation of person B has level X_2 .

then after duration Δt the sensory representation of person B will have level $X_2 + \eta [c(\omega_{1B}X_1) - X_2] \Delta t$.

$$d \text{ srs}_B(t)/dt = \eta [c(\omega_{1B} \text{ srs}_B(t)) - \text{ srs}_B(t)]$$

Here $c(\dots)$ is a combination function for which different choices can be made, for example, the identity function $\mathbf{id}(V) = V$, or an advanced logistic sum function of the form

$$\mathbf{alogistic}_{\sigma, \tau}(V_1, \dots, V_k) = \left(\frac{1}{1 + e^{-\sigma(V_1 + \dots + V_k - \tau)}} - \frac{1}{1 + e^{\sigma\tau}} \right) (1 + e^{-\sigma\tau})$$

with σ a steepness and τ a threshold value. In the simulations for the sake of simplicity for properties LP1, LP2, and LP11–LP14 the identity combination function $c(V) = \mathbf{id}(V) = V$ was chosen for $c(\dots)$. For properties LP3–LP10 the logistic sum combination function was used. Property LP2 is similar to LP1 but applied to stimulus s instead of person B (for example, a face expression).

LP2 Sensory representation of stimulus s

If a stimulus s is sensed with level X_1 ,

and the sensory representation of s has level X_2 .

then after duration Δt the sensory representation of s will have level $X_2 + \eta [c(\omega_2 X_1) - X_2] \Delta t$.

$$\mathbf{d} \text{ srs}_s(t)/\mathbf{d}t = \eta [c(\omega_2 \text{ srs}_s(t)) - \text{srs}_s(t)]$$

The sensory representation of a body state as described by property LP3 is not only affected by a corresponding sensor state (which in turn is affected by the body loop), but also via the as-if body loop by the preparation for this body state. Note that the as-if body loop provides effects on the sensory representation in a shorter time than via the body loop: bodily change usually is a factor slower than neurological change (e.g., 1 or 2 s vs. 300–500 ms).

LP3 Sensory representation of a body state

If the sensor state for body state b has level X_1

and the preparation state for body state b has level X_2

and the sensory representation of body state b has level X_3

then after Δt the sensory representation of body state b will have level $X_3 + \eta [c(\omega_3 X_1, \omega_4 X_2) - X_3] \Delta t$.

$$\mathbf{d} \text{ srs}_b(t)/\mathbf{d}t = \eta [c(\omega_3 \text{ srs}_b(t), \omega_4 \text{ ps}_b(t)) - \text{srs}_b(t)]$$

Generating preparation, mirroring and control states

Preparation for a bodily change triggered by s (e.g., an observed face expression leading to preparation for a similar expression) is modeled as follows.

LP4 Preparing for or mirroring a body state

If the sensory representation of s has level X_1 ,

and the sensory representation of b has level X_2 ,

and the preparation for body state b has level X_3

then after duration Δt the preparation state for body state b will have level $X_3 + \eta [c(\omega_5 X_1, \omega_6 X_2) - X_3] \Delta t$.

$$\mathbf{d} \text{ ps}_b(t)/\mathbf{d}t = \eta [c(\omega_5 \text{ srs}_s(t), \omega_6 \text{ srs}_b(t)) - \text{ps}_b(t)]$$

Control for a person B generates a state indicating on which person (self-other distinction for B another person) the focus is, and whether or not to act (the case of self); this is modeled in LP5.

LP5 Control for another person or self

If the sensory representation of person B (another person or self) has level X_1 ,

and the sensory representation of s has level X_2 ,

and the control state for B , s and b has level X_3

then after duration Δt the control state for B , s and b will have level $X_3 + \eta [c(\omega_{7B} X_1, \omega_{8B} X_2) - X_3] \Delta t$.

$$\mathbf{d} \text{ cs}_{B,s,b}(t)/\mathbf{d}t = \eta [c(\omega_{7B} \text{ srs}_B(t), \omega_{8B} \text{ srs}_s(t)) - \text{cs}_{B,s,b}(t)]$$

Control for sensory processing sensitivity, modeled in LP6, generates a state indicating in how far the stimulus induces an inadequately high sensory body representation level. This state is the basis for two possible regulations (modeled in LP8 and LP10 below): of the expressed body state, and of the gaze.

LP6 Control for enhanced sensitivity

If the sensory representation of s has level X_1 ,
 and the sensory representation of b has level X_2
 and the sensitivity control state for s and b has level X_3
 then after duration Δt the sensitivity control state for s and b will have level $X_3 + \eta [c(\omega_9 X_1, \omega_{10} X_2) - X_3] \Delta t$.

$$\mathbf{d} \, cs_{sens,s,b}(t)/\mathbf{d}t = \eta [c(\omega_9 \, srs_s(t), \omega_{10} \, srs_b(t)) - cs_{sens,s,b}(t)]$$

The preparation of a verbal empathic reaction to another person depends on feeling a similar emotion, and on adequate self-other distinction, as modeled in LP7.

LP7 Preparing for communication

If the sensory representation of body state b has level X_1 ,
 and the control for person $B \neq \text{self}$, s and b has level X_2 ,
 and the preparation of communication of b to B has level X_3
 then after Δt the preparation of communication of b to B will have level $X_3 + \eta [c(\omega_{11B} X_1, \omega_{12B} X_2) - X_3] \Delta t$.

$$\mathbf{d} \, ps_{B,b}(t)/\mathbf{d}t = \eta [c(\omega_{11B} \, srs_b(t), \omega_{12B} \, cs_{B,s,b}(t)) - ps_{B,b}(t)]$$

Expressing Prepared States

Expressing a (prepared) body state depends on whether a control state for self is available. However, to cover regulative behaviour to compensate for enhanced sensory processing sensitivity, also the sensitivity control state is involved, with an inhibiting effect on expressing the prepared body state (ω_{14} is taken to be negative). Such an effect can achieve that although the person feels the same as the other person, the face remains expressionless. In this way LP8 models a mechanism for *response-focused regulation* (suppression of the own response) to compensate for an undesired level of arousal; e.g., Gross (1998), Goldin et al. (2008).

LP8 Expressing a body state

If the control state for self, s and b has level X_1 ,
 and the control state for sensitivity, s and b has level X_2 ,
 and the preparation for body state b has level X_3
 and expressing body state b has level X_4
 then after duration Δt body state b will be expressed with level $X_4 + \eta [c(\omega_{13} X_1, \omega_{14} X_2, \omega_{15} X_3) - X_4] \Delta t$.

$$\mathbf{d} \, es_b(t)/\mathbf{d}t = \eta [c(\omega_{13} \, ps_{self,s,b}(t), \omega_{14} \, cs_{sens,s,b}(t), \omega_{15} \, ps_b(t)) - es_b(t)]$$

Note that expression states es are the person's execution states (e.g., the muscle states); body and gaze states result from these expression states (via LP11 and LP12 below). A preparation for a verbal empathic reaction leads to expressing this communication in a straightforward manner.

LP9 Expressing communication

If the preparation of communication of b to B has level X_1 ,
 and the expressed communication for b to B has level X_2
 then after Δt the person will express communication of b to B with level $X_2 + \eta [c(\omega_{16B} X_1) - X_2] \Delta t$.

$$\mathbf{d} \text{esc}_{B,b}(t)/\mathbf{d}t = \eta [c(\omega_{16B} \text{ps}_{B,b}(t)) - \text{esc}_{B,b}(t)]$$

Dynamic property LP10 models *antecedent-focused regulation (attentional deployment)* as described in Gross (1998), Goldin et al. (2008): directing the own gaze away from the stimulus that feels too overwhelming. Note that the gaze direction avoiding s is chosen to be 1 for total avoidance of stimulus s , and 0 for no avoidance (it indicates the extent of avoidance).

LP10 Expressing gaze for avoidance of s

If the control state for sensitivity, s and b has level X_1 ,
and the expressed gaze for avoidance of s has level X_2
then after Δt the expressed gaze for avoidance of s will have level $X_2 + \eta [c(\omega_{17}X_1) - X_2] \Delta t$.

$$\mathbf{d} \text{es}_{\text{avoiding},s}(t)/\mathbf{d}t = \eta [c(\omega_{17} \text{cs}_{\text{sens } s,b}(t)) - \text{es}_{\text{avoiding},s}(t)]$$

Maintaining body and gaze states

Properties LP11 and LP12 describe how the expression states affect the body and gaze in a straightforward manner.

LP11 From body expression to body state

If the expression state for body state b has level X_1 ,
and the body state b has level X_2
then after Δt body state b will have level $X_2 + \eta [c(\omega_{18}X_1) - X_2] \Delta t$.

$$\mathbf{d} \text{ws}_b(t)/\mathbf{d}t = \eta [c(\omega_{18} \text{es}_b(t)) - \text{ws}_b(t)]$$

LP12 is similar to LP11 with gaze instead of body.

LP12 From gaze avoidance expression to gaze avoidance state

If the expression state for gaze avoidance for s has level X_1 ,
and the gaze avoidance for s has level X_2
then after Δt the gaze avoidance for s will have level $X_2 + \eta [c(\omega_{19}X_1) - X_2] \Delta t$.

$$\mathbf{d} \text{ws}_{\text{avoiding},s}(t)/\mathbf{d}t = \eta [c(\omega_{19} \text{es}_{\text{avoiding},s}(t)) - \text{ws}_{\text{avoiding},s}(t)]$$

Generating sensor states

Sensing a body state and person B also happen in a straightforward manner, as described by LP13 and LP14.

LP13 Generating a sensor state for a body state

If the body state b has level X_1 ,
and the sensor state for body state b has level X_2
then after Δt the sensor state for body state b will have level $X_2 + \eta [c(\omega_{20}X_1) - X_2] \Delta t$

$$\mathbf{d} \text{ss}_b(t)/\mathbf{d}t = \eta [c(\omega_{20} \text{ws}_b(t)) - \text{ss}_b(t)]$$

LP14 is similar to LP13 with person B instead of body.

LP14 Generating a sensor state for person B

If person B is present with level X_1 ,
and the sensor state for person B has level X_2

then after Δt the sensor state for person B will have level $X_2 + \eta [c(\omega_{21}X_1) - X_2] \Delta t$
 $\mathbf{d} \text{ ss}_B(t)/\mathbf{d}t = \eta [c(\omega_{21} \text{ ws}_B(t)) - \text{ss}_B(t)]$

Within the external world, to generate a sensor state for a stimulus s , the gaze state with respect to s is taken into account. As the gaze state indicates the extent of avoidance of s , it has an inhibiting effect on sensing s (ω_{23} is chosen negative); here the combination function $c(\cdot)$ has been modeled by $c(V_1, V_2) = V_1(1 + V_2)$ with $-1 \leq V_2 \leq 0$. Note that this is an example of a combination function which is not symmetric. Here the first argument V_1 refers to the impact from the connection from ws_s , and the second argument V_2 refers to the impact by the connection from $\text{ws}_{\text{avoiding},s}$. This makes that for $V_2 = -1$ it results in $c(V_1, V_2) = 0$, and for $V_2 = 0$ it results in $c(V_1, V_2) = V_1$.

LP15 Generating a sensor state for a stimulus

If stimulus s is present with level X_1 ,

and gaze avoidance state for s has level X_2 ,

and the sensor state for s has level X_3 ,

then after Δt the sensor state for s will have level $X_3 + \eta [c(\omega_{22}X_1, \omega_{23}X_2) - X_3] \Delta t$
 $\mathbf{d} \text{ ss}_s(t)/\mathbf{d}t = \eta [c(\omega_{22} \text{ ws}_s(t), \omega_{23} \text{ ws}_{\text{avoiding},s}(t)) - \text{ss}_s(t)]$

LP1 Sensory representation of an external agent B

$\mathbf{d} \text{ srs}_B(t) / \mathbf{d}t = \eta [c(\omega_{1B} \text{ ss}_B(t)) - \text{srs}_B(t)]$

LP2 Sensory representation of a stimulus s

$\mathbf{d} \text{ srs}_s(t) / \mathbf{d}t = \eta [c(\omega_2 \text{ ss}_s(t)) - \text{srs}_s(t)]$

LP3 Sensory representation of a body state

$\mathbf{d} \text{ srs}_b(t) / \mathbf{d}t = \eta [c(\omega_3 \text{ ss}_b(t), \omega_4 \text{ ps}_b(t)) - \text{srs}_b(t)]$

LP4 Preparing for or mirroring a body state

$\mathbf{d} \text{ ps}_b(t) / \mathbf{d}t = \eta [c(\omega_5$
 $\text{srs}_s(t), \omega_6 \text{ srs}_b(t)) - \text{ps}_b(t)]$

LP5 Control for another agent or self

$\mathbf{d} \text{ cs}_{B,s,b}(t) / \mathbf{d}t = \eta [c(\omega_{7B} \text{ srs}_B(t), \omega_{8B} \text{ srs}_s(t)) - \text{cs}_{B,s,b}(t)]$

LP6 Control for enhanced sensitivity

$\mathbf{d} \text{ cs}_{\text{sens},s,b}(t) / \mathbf{d}t = \eta [c(\omega_9 \text{ srs}_s(t), \omega_{10} \text{ srs}_b(t)) - \text{cs}_{\text{sens},s,b}(t)]$

LP7 Preparing for communication

$\mathbf{d} \text{ ps}_{B,b}(t) / \mathbf{d}t = \eta [c(\omega_{11B} \text{ srs}_b(t), \omega_{12B} \text{ cs}_{B,s,b}(t)) - \text{ps}_{B,b}(t)]$

LP8 Expressing a body state

$\mathbf{d} \text{ es}_b(t) / \mathbf{d}t = \eta [c(\omega_{13} \text{ ps}_{\text{self},s,b}(t), \omega_{14} \text{ cs}_{\text{sens},s,b}(t), \omega_{15} \text{ ps}_b(t)) - \text{es}_b(t)]$

LP9 Expressing communication

$\mathbf{d} \text{ esc}_{B,b}(t) / \mathbf{d}t = \eta [c(\omega_{16B} \text{ ps}_{B,b}(t)) - \text{esc}_{B,b}(t)]$

LP10 Expressing gaze for avoidance of s

$$d \text{ es}_{\text{avoiding},s}(t) / dt = \eta [c(\omega_{17} \text{ cs}_{\text{sens } s,b}(t)) - \text{es}_{\text{avoiding},s}(t)]$$

LP11 From body expression to body state

$$d \text{ ws}_b(t) / dt = \eta [c(\omega_{18} \text{ es}_b(t)) - \text{ws}_b(t)]$$

LP12 From gaze expression to gaze state

$$d \text{ ws}_{\text{avoiding},s}(t) / dt = \eta [c(\omega_{19} \text{ es}_{\text{avoiding},s}(t)) - \text{ws}_{\text{avoiding},s}(t)]$$

LP13 Generating a sensor state for a body state b

$$d \text{ ss}_b(t) / dt = \eta [c(\omega_{20} \text{ ws}_b(t)) - \text{ss}_b(t)]$$

LP14 Generating a sensor state for an agent B

$$d \text{ ss}_B(t) / dt = \eta [c(\omega_{21} \text{ ws}_B(t)) - \text{ss}_B(t)]$$

LP15 Generating a sensor state for a stimulus

$$d \text{ ss}_s(t) / dt = \eta [c(\omega_{22} \text{ ws}_s(t), \omega_{23} \text{ ws}_{\text{avoiding},s}(t)) - \text{ss}_s(t)]$$

Box 9.1 Overview of the numerical representation of the temporal-causal network model

9.4 Types of Social Response Patterns Shown

To analyse the different types of response patterns shown by the temporal-causal network model, some dynamic properties were identified. By automated verification (see Chap. 13) they have been checked for generated simulation traces, allowing to evaluate easily the patterns for a variety of parameter values. Below the dynamic properties are introduced in an informally expressed in a conceptual manner; their formalisations by numerical-logical expressions are shown in Chap. 13.

9.4.1 Overview of Basic Patterns

The simulations discussed first, have been performed with $\eta = 1$, $\Delta t = 0.5$, and settings for threshold and steepness values as shown in Table 9.2. In the graphs in Figs. 9.2 and 9.3, and further, time is at the horizontal axis and activation levels are at the vertical axis.

Table 9.2 Setting for threshold and steepness values used

	LP3	LP4	LP5self	LP5sens	LP6	LP7	LP8	LP9	LP10
τ	0.8	1	1	1	2.5	1.5	1.5	0.5	0.5
σ	8	8	40	40	40	8	40	40	40

Fig. 9.2 Example simulations: full empathic response

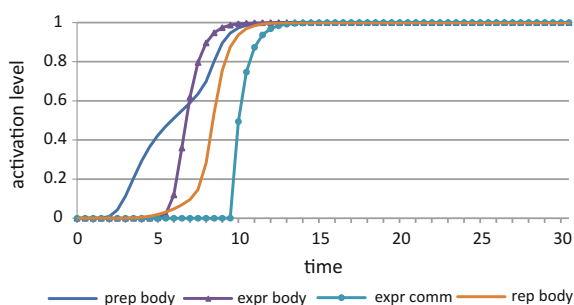
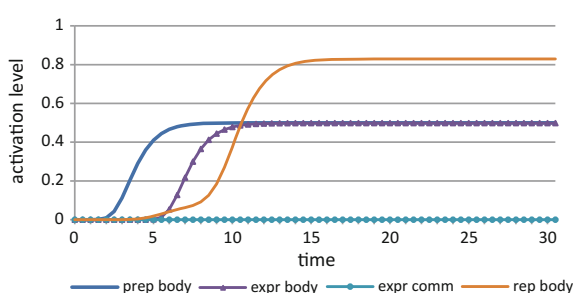


Fig. 9.3 Example simulations: reduced emotion integration



The first property expresses that when a person B is met, showing a certain emotion, then within a certain time a response occurs, which can consist of:

- (1) person self feels the same as person B ,
- (2) this feeling is bodily expressed by self, and
- (3) it is communicated by person self to person B that B feels this.

An example of this is the response described for Example 1 in Sect. 9.2.5.

SBP1 Response occurrence

When person $B \neq$ self is present expressing a certain feeling b from some point in time on, then after some time person self will have a response R (generating the feeling of b , resp. bodily expression, resp. communication).

By combination 8 different types of response are possible; see Table 9.3. Some of them are not likely to occur (types 5, 6, and 7): when the person self does not feel the emotion, it is probably hard to communicate or show it. The way in which different connections relate to different types of processes, as depicted in Table 9.1, provides an indication of which deviant connection strengths may lead to which phenomena. For example, when ω_5 (connecting srs_s to ps_b ; see Fig. 9.1 and Table 9.1) is low, mirroring is reduced, and as a consequence a low social response (type 8) occurs, which is in accordance with what is reported, for example, in Dapretto et al. (2006). This may correspond to the missed empathic response in Sect. 9.2.5, Example 2.

An example of type 1 is shown in Fig. 9.2 displaying the feeling state srs_b (rep body), preparation due to mirroring ps_b (prep body), expression of body by

Table 9.3 Different types of possible social responses: + means that the response occurs, – means that it does not occur (or is very weak)

	1	2	3	4	5	6	7	8
Feeling	+	+	+	+	–	–	–	–
Body	+	+	–	–	+	+	–	–
Communication	+	–	+	–	+	–	+	–
Type of response	Full empathic response	Feeling and body expression without communication	Feeling and communication without body expression	Feeling without body expression and communication				No feeling, no body expression, no communication
Example conditions	Fully adequate conditions	Inadequate self-other distinction	Inadequate emotion integration	Both inadequate emotion integration and self-other distinction				Inadequate mirroring
Example parameter settings	None of ω_k low	ω_{7B} , ω_{8B} low or ω_{11B} , ω_{12B} low	$\omega_{7\text{self}}$, $\omega_{8\text{self}}$ low	ω_6 , ω_{11} low				ω_5 low

execution state es_b (expr body), and communication state $esco_{B,b}$ (expr comm). Here, $\omega_k = 1$ for all k , except for the suppressing connections (from $cs_{sens,s,b}$ to es_b , and from $ws_{avoiding,s}$ to ss_s , respectively): $\omega_{14} = \omega_{23} = -1$. The pattern shows an increase of mirroring, followed by bodily expression and feeling, and communication. This corresponds to a type of response as shown by the physician in Sect. 9.2.5, Example 1.

Response type 4 in Table 9.3 only concerns the feeling (not externally observable). For response type 2, the feeling is expressed: it is externally observable, but no verbal communication takes place. Response type 2 with low ω_{7B} or ω_{8B} (from srs_B , resp. srs_s to $cs_{B,s,b}$) displays that no adequate self-other distinction is made (reduced control); e.g., Iacoboni (2008a), Brass and Spengler (2009).

Response type 4 with low ω_6 (from srs_b to ps_b) can be viewed as a form of emotion contagion without integrating the emotion in responses; e.g., Grèzes and de Gelder (2009), Grèzes et al. (2009). In contrast, in response type 3 the emotion felt is attributed to the other person, but no bodily expression is shown. Figure 9.3 shows an example of response type 4. The level of emotion felt is becoming high, but due to lack of emotion integration ($\omega_6 = \omega_{11B} = 0$ and the other ω_k the same as for the graph in Fig. 9.2), the bodily and verbal expression are reduced.

In case of regulation due to enhanced sensory sensitivity (e.g., Baker et al. 2008; Corden et al. 2008), patterns occur when a response only lasts for a short time, expressed as:

SBP2 Response withdrawal

When person $B \neq \text{self}$ is present expressing a certain feeling b from some point in time on, and the person self has response R , then within time duration D this response will disappear.

9.4.2 Oscillatory Patterns: Limit Cycle Behaviour

The combination SBP1 and not SBP2 expresses a persistent response, whereas SBP1 and SBP2 specifies only a short occurrence of a response. However, after withdrawal of the response due to regulation, also the arousal level for b will become low, which brings the person in practically the same state as initially. An oscillatory pattern results, while the environment is fully static. Such oscillatory social response patterns indeed can be observed in persons with some forms of ASD, who let their gaze go back and forth to another person's eyes during a contact, as a way of regulation of enhanced sensitivity. Figure 9.4 shows an example of such a response pattern, specified as follows.

SBP3 Response oscillation

When a person B bodily expressing a certain feeling is present from some point in time on, then:

- (1) for every time point there is a later time point for which response R occurs
- (2) for every time point there is a later time point for which response R does not occur

The model shows this type of social response when the threshold for sensory sensitivity is set between 1 and 2; for example, for Fig. 9.4, it was set to 1.2. Moreover, as for the upper graph $\omega_k = 1$ for all k , except for the suppressing connections: $\omega_{14} = \omega_{23} = -1$. It is shown that body expression and communication

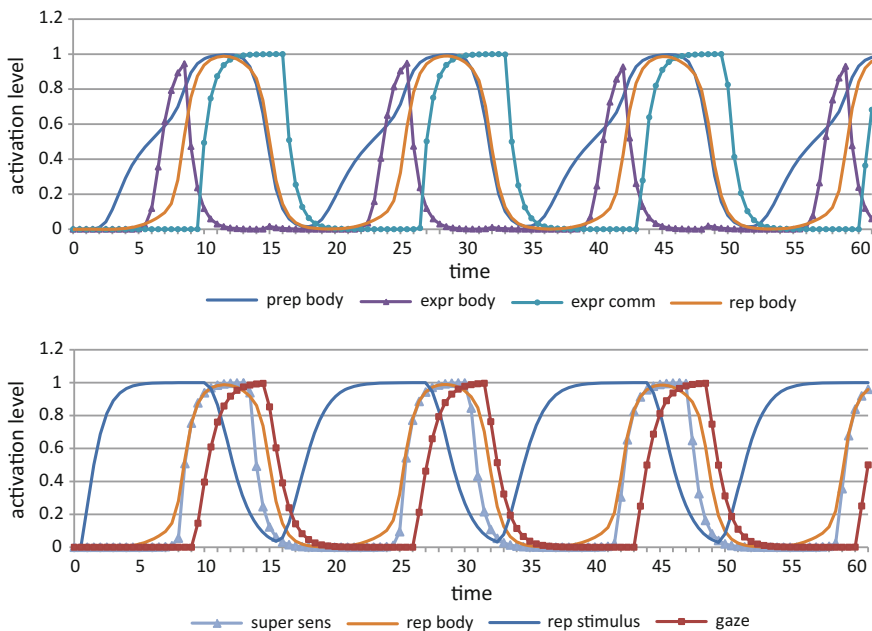


Fig. 9.4 Example simulation: enhanced sensitivity

last only for short time periods, but recur. If the threshold value is set 1 or lower, no response occurs (type 8); if it is 2 or higher a persistent response occurs (type 1).

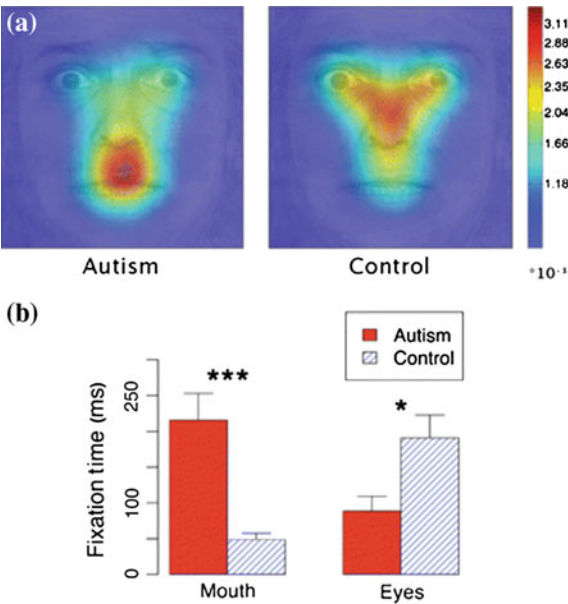
Note that instead of varying the threshold for sensory sensitivity, similar patterns are generated when the connection strength ω_{17} (from $cs_{sens,s,b}$ to $es_{avoiding,s}$) is varied. The oscillatory patterns due to regulation for enhanced sensitivity occur for all response types in Table 9.3.

Note that the oscillatory patterns shown by the model have an inherent regular periodicity, even without any alternating world events offered to the model: the environment is kept constant. For example, in Fig. 9.4 the pattern repeats itself about every 18 time units. This period reflects the time needed to calm down the too high arousal. This type of behaviour of the model is an example of a reaching a *limit cycle*; for more details about this type of behaviour, see Chap. 12, Sect. 12.7.

9.4.3 Comparison to Empirical Gaze Data

From human experiments empirical data of gaze patterns are known. These can be used to have a (quite modest) validation of the model. For example, in Neumann et al. (2006) it was found that for a group of persons with ASD the average fixation time of the gaze was about 85 ms at the eyes and 215 ms at the mouth (for the control group this was 190 ms at the eyes vs. 50 ms at the mouth); e.g., Neumann et al. (2006, p. 198, Fig. 4), see also Fig. 9.5.

Fig. 9.5 Fixations made in an experiment, adopted from Neumann et al. (2006, Fig. 4)



This gives an estimated average period of 300 ms for the ASD group in that experiment. It was possible to mimic this average period quite accurately by choosing $\eta = 0.06$ and keeping all other parameter value the same as for the case depicted in Fig. 9.4. So, for these settings the model describes an average of this group of persons with ASD. Note, however, that individuals will deviate from this average, and one individual will also show differences over time. In realistic situations eye movements do not only depend on the regulation mechanism, but also on other events that may happen, for example, a gesture or an eye-blink of the other person, or some inter-punctuation in the talking of the other person.

9.4.4 Interaction of Two Persons Displaying Regulation of Enhanced Sensory Sensitivity

In the scenarios discussed above and shown in Figs. 9.2, 9.3 and 9.4 the other person B and the stimulus were assumed static. However, the model can be applied to person B as well. In this case it is assumed that the eyes of one person are the stimulus for the other person, so that in a mutual manner an avoiding gaze regulation of one person affects the stimulus for the other person as well. This might describe a situation when a physician with reduced social interaction capabilities (due to enhanced sensory processing sensitivity) has an interaction with a patient who also has such reduced capabilities. For this situation, it turns out that the interaction often starts in an asynchronous and irregular way, as shown in Fig. 9.6. This is an example where the values for one parameter, namely the update speed parameter η were taken different. This can be considered as expressing an individual difference in neurological response time: for person A it is 1 as for self before, and for person B it is 0.7, which means that person B responds 30 % slower than person A . Also when differences in values of other parameters or in initial values are made, the pattern starts in an asynchronous and irregular manner. In Fig. 9.6 the upper half of the figure shows activation levels of person A over time, the lower half the same for person B . For example, in Fig. 9.6, first both gazes are on the eyes, but after time point 10 person A takes the gaze away from the eyes (until time point 17). Person B also starts to take away the gaze from the eyes at time 15, but soon comes back again (around time 20), since in the meantime the other person's gaze has gone elsewhere. But after time 20 person A 's gaze comes on the eyes again and then person B takes the gaze from the eyes for a longer time (until time 30). Such social interaction patterns may occur as a bit chaotic and weird (the pattern shown will in fact later on end up in a periodic oscillating pattern: a limit cycle), but the introduced temporal-causal network model shows the logic and rationality behind such patterns.

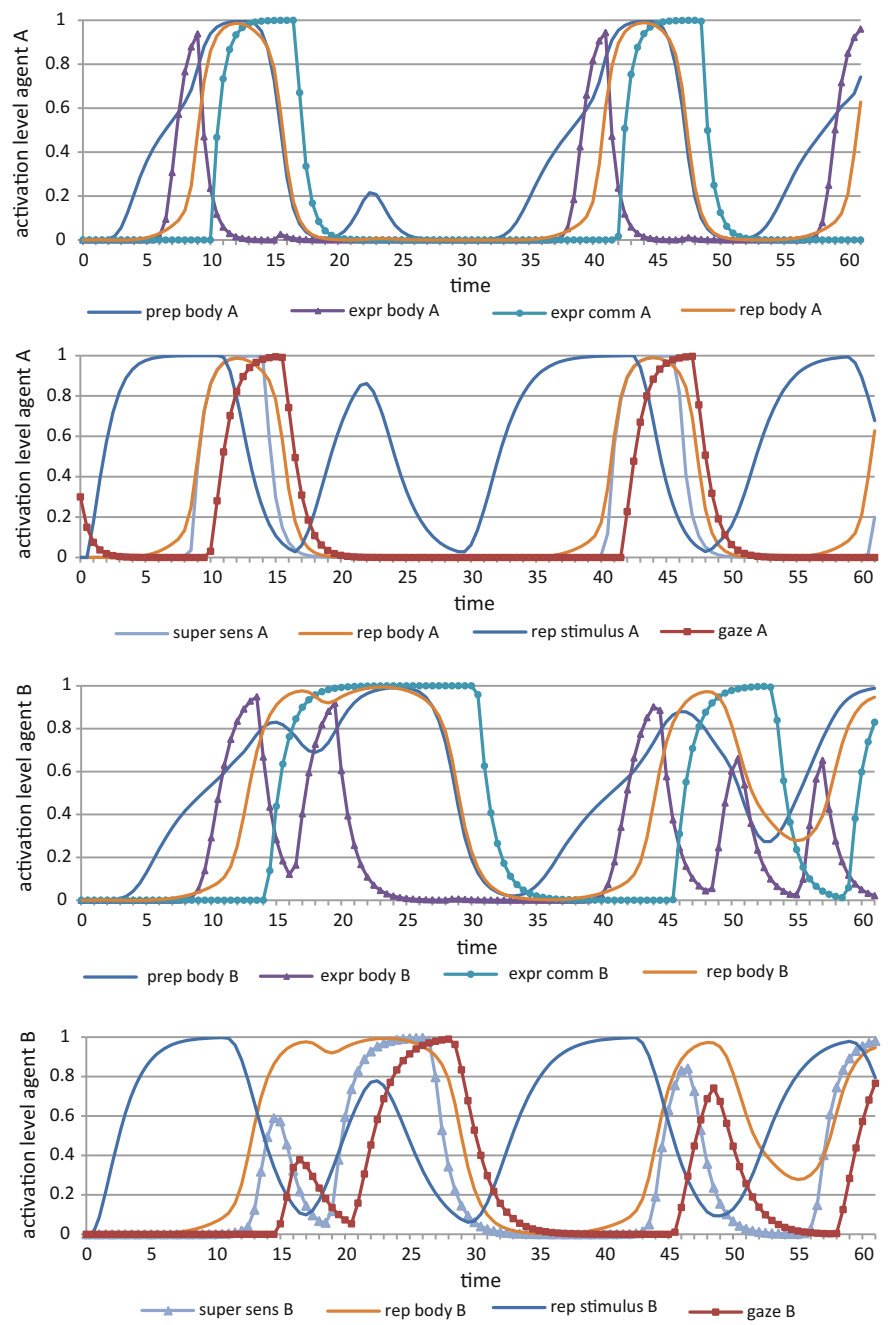


Fig. 9.6 Bidirectional enhanced sensitivity pattern

$$\frac{d\omega_{i,j}(t)}{dt} = \eta X_i(t)X_j(t)(1 - \omega_{i,j}(t)) - \zeta \omega_{i,j}(t)$$

$$\omega_{i,j}(t + \Delta t) = \omega_{i,j}(t) + [\eta X_i(t)X_j(t)(1 - \omega_{i,j}(t)) - \zeta \omega_{i,j}(t)]\Delta t \quad (9.1)$$

A similar Hebbian learning rule can be found in Gerstner and Kistler (2002, p. 406). By the factor $1 - \omega_{i,j}(t)$ the learning rule keeps the level of $\omega_{i,j}(t)$ bounded by 1 (which could be replaced by any other positive number); Hebbian learning without such a bound usually provides instability. When the extinction rate is relatively low, the upward changes during learning are proportional to both $X_i(t)$ and $X_j(t)$ and maximal learning takes place when both are 1. Whenever one of them is 0 (or close to 0) extinction takes over, and ω slowly decreases (unlearning). This learning principle has been applied (simultaneously) to all six connections indicated by dotted arrows in Fig. 9.1; for the sake of transparency, the 12 arrows to the connection weights have been left out. In principle, the learning rate η and extinction rate ζ , could be taken differently for the different dynamical relationships. In the example simulations discussed in Sect. 9.6 uniform values have been used: $\eta = 0.16$ and $\zeta = 0.0032$.

Most of the connections have been given strength 1 or -1 , but six of them (indicated by dotted arrows) have a dynamical strength, adapted over time according to Hebbian learning.

9.6 Example Simulations of Learning Processes

A number of simulation experiments have been conducted for different types of scenarios, using numerical software. For the examples discussed here the values for the threshold and steepness parameters are as shown in Table 9.4. Note that first the value 3 for sensitivity control threshold was chosen so high that no enhanced sensitivity occurs. The learning rate was set to $\eta = 0.16$ and extinction rate

Table 9.4 Settings for threshold and steepness parameters

		τ	σ
Representing body state	srs_b	1	3
Control B	$cs_{B,s,b}$	0.7	30
Control sensitivity	$cs_{sens,s,b}$	3	30
Mirroring/preparing body state	ps_b	1	3
Preparing communication	$ps_{b,B}$	0.8	3
Expressing body state	es_b	1.2	30
Expressing communication	$esc_{b,B}$	0.8	30
Expressing gaze avoidance state	$es_{avoiding,s}$	0.6	30

$\zeta = 0.0032$. The step size Δt was set to 1. All nonadapted connection strengths have been given value 1, except those for suppressing connections

$$\omega_{cs_{sens,s,b}ps_b}, \omega_{cs_{sens,s,b}es_b} \text{ and } \omega_{ws_{avoiding,s,ss_s}}$$

which have been given the value -1 . The scenario was chosen in such a way that after every 100 time units another person is encountered for a time duration of 25 units with a body expression that serves as stimulus. Initial values for activation levels of the internal states were set at 0. A first pattern, displayed in Fig. 9.8, is that in normal circumstances, assuming initial strengths of the learned connections of 0.3, the model is indeed able to learn the empathic responses as expected. Here (and also in Fig. 9.9) time is on the horizontal axis and activation levels at the vertical axis.

The upper graph shows levels for body representation, body preparation, expressed body states and communication. The lower graph shows the learning

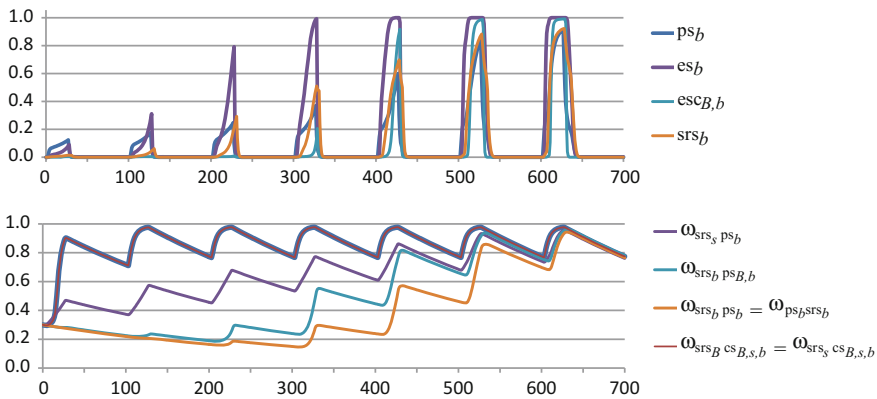


Fig. 9.8 Example scenario of the Hebbian learning process

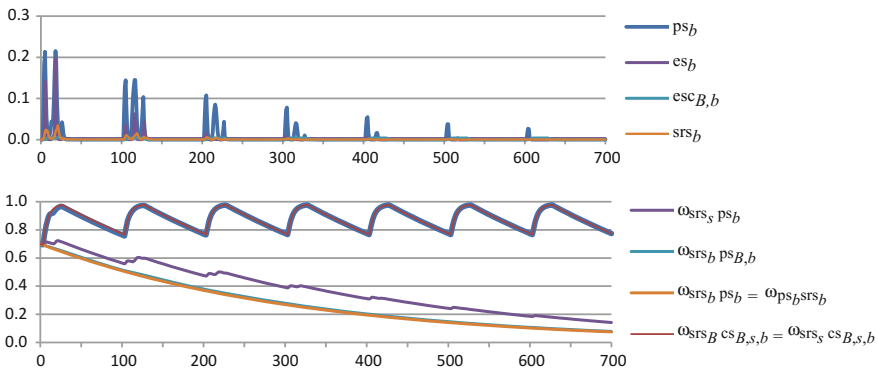


Fig. 9.9 Learning under enhanced sensory processing sensitivity

patterns for the connections (the dotted arrows in Fig. 9.7). Note that the two connections

ω_{srs_b, ps_b} (for emotion integration) and ω_{ps_b, srs_b} (as-if body loop)

turn out to have the same values, as they connect the same states srs_b and ps_b , and have been given the same initial values. Moreover, also the connections

$\omega_{srs_B, cs_{B,s,b}}$ and $\omega_{srs_s, cs_{B,s,b}}$

have the same values, as in the considered scenario the input states for srs_B and srs_s have been given the same values, and also the initial values for the connections. This can easily be varied. In Fig. 9.8 it is shown that when regular social encounters take place, the connections involved in responding empathically are strengthened to values that approximate 1. Notice that due to the relatively low initial values of the connections chosen, for some of them first extinction dominates, but later on this downward trend is changing into an upward trend. Accordingly the empathic responses become much stronger, which is in line with the literature; e.g., Field et al. (2001), Ingersoll et al. (2007).

How long the learned patterns will last will depend on the social context. When after learning the person is isolated from any social contact, the learned social behaviours may vanish due to extinction. However, if a certain extent of social contact is offered from time to time, the learned behaviour is maintained well. This illustrates the importance of the social context. When zero or very low initial levels for the connections are given, this natural learning process does not work. However, as other simulations show, in such a case (simulated) imitation training sessions (starting with the therapist imitating the person) still have a positive effect, which is also lasting when an appropriate social context is available. This is confirmed by reports that imitation training sessions are successful; e.g., Field et al. (2001), Ingersoll et al. (2007).

In addition to variations in social environment, circumstances may differ in other respects as well. From many persons with some form of autistic spectrum disorder it is known that they show enhanced sensory processing sensitivity; e.g., Baker et al. (2008), Corden et al. (2008); this was also incorporated in the model. Due to this, their regulation mechanisms to avoid a too high level of arousal may interfere with the social behaviour and the learning processes. Indeed, in simulation scenarios for this case it is shown that the adaptive model shows an unlearning process: connection levels become lower instead of higher. This pattern is shown in Fig. 9.9. Here the same settings are used as in Table 9.1, except the sensitivity control threshold which was set at 1 in this case, and the initial values for the connection weights, which were taken 0.7. It is shown that the connections

ω_{srs_s, ps_b} (for mirror activation) and ω_{srs_b, ps_b} and ω_{ps_b, srs_b} (for emotion integration)

are decreasing, so that the responses become lower over time.

This is due to the downregulation which, for example, leads to a gaze that after a short time is taken away from the stimulus, and returns after the arousal has decreased, after which the same pattern is repeated; this is shown in the upper graph

(the two or three peaks per encounter). Note that the values related to control of and communication to another person stay high: the downregulation as modeled does not have a direct effect on these processes. When downregulation is also applied to communication, also these connections will extinguish. When for such a case imitation training sessions are offered in a simulation, still the connection levels may be strengthened. However, these effects may not last in the natural context: as soon as these sessions finish, the natural processes may start to undo the learned effects. To maintain the learned effects for this case such training sessions may have to be repeated regularly.

9.7 Discussion

The presented temporal-causal network model for regulated social response patterns uses theories from Social Neuroscience as a point of departure: theories on mirror neuron systems, self-other distinction, emotion integration, emotion regulation, and empathy. The contents of this chapter are based on Treur (2011a, b, 2014). It was shown how a wide variety of realistic social response patterns can be obtained by varying the person's makeup of mental structures, inspired by relevant literature on autism spectrum disorders. In contrast to work as discussed in Hendriks and Treur (2010), Laan and Treur (2011), Bosse et al. (2012), the presented temporal-causal network model addresses regulation of enhanced sensory processing sensitivity by control of body, face expression and gaze, based on the emotion regulation theory presented in Gross (1998, 2007), Goldin et al. (2008). This model was extended by Hebbian learning, thus obtaining an adaptive temporal-causal network model. It was shown how in this adaptive network model learning of empathic social interaction takes place and also how enhanced sensory processing sensitivity reduces these learning effects. A modest step towards validation that has been taken has been discussed. The model has been used as a point of departure to model internalizing and externalizing behavior for persons with ASD; see Lubbe et al. (2016).

The temporal-causal network model provides a basis for human-like behaviour of virtual agents in the context of simulation-based training or gaming. For example, it may provide a basis for the implementation of virtual agents for training of professionals such as teachers, psychotherapists or physicians, or in applications of human-like virtual characters with realistic body and face expression and gaze. For example, in Suchman (1997), Tulskey (2011) the need for such training of physicians and computer assistance for such training is emphasized. The presented model provides a basis to easily generate different types of example interaction scenarios varying from perfect ones to instances with some aspect(s) of imperfection, and use these to deepen insight in such responses and their possible imperfections. For example, an environment can be developed offering a virtual agent which is able to show face expressions and communication, which is driven by the temporal-causal network model presented here. Interaction with a second

virtual agent can be shown within this environment. Such an environment can have possibilities to easily adapt parameters (e.g., the ω 's) and see what the effect is on the social interaction between the two virtual agents. This environment can be used as a tool during the education of psychotherapists, to obtain (virtual) experiences with different forms and deviations in social interaction in relation to these parameter settings.

In a wide literature, the role of emotions in virtual agents in general is addressed; e.g., Bates et al. (1994), Yang et al. (2008), Gratch et al. (2009). Usually these approaches are not specifically related to empathic responses, and often use body or face expressions as a way of presentation, and not as a more biologically grounded basis for the emotion as in the neurological perspective of Damasio (1999), which was adopted in the current paper. The importance of computational models for 'caring' agents in a virtual context showing empathy has also been recognized in the literature; see, for example Klein et al. (2002), Bickmore and Picard (2004), McQuiggan et al. (2008), Bickmore et al. (2010). Moreover, in Ochs et al. (2008), Rodrigues et al. (2009), Boukricha and Wachsmuth (2011), Paiva (2011), Leite et al. (2012) virtual agents are developed that have or show empathy. In this literature the aim is to realize perfect empathy. The basis is usually chosen in appraisal theories for emotion generation. The temporal-causal network model presented in the current paper differs from such existing models in that it is grounded in recent insights from neuroscience and emotion regulation, and reflects these theories. Moreover, the presented model is able to display social responses in a realistic human-like manner, not only of ideal empathic humans, but also of socially less perfect humans. Therefore using the current model it is possible, for example, in simulation-based training to generate example scenarios showing certain forms of imperfection in social interaction which are realistic in the sense that they directly relate to differences in the human population as described by the neurological theories used as a basis.

In McQuiggan et al. (2008) the CARE framework for experiments with humans and empathic virtual agents is described. A possibility for future research is to integrate the presented model in an environment such as, for example the CARE environment, and conduct experiments with different types of (imperfect) empathic agents. As another example, based on the presented model a social interaction pattern between two persons as shown in Fig. 9.3 can be easily implemented within a displayed virtual story context. The expressed emotions can be displayed on the faces of the two agents, and gaze regulation can be displayed as eyes or faces turning away from each other. When the model described is used as an engine to generate the states and behaviour for each of the two virtual agents, the interactive pattern will automatically be generated.

Modeling causal relations discussed in neurological literature in the manner as presented here does not need to take specific neurons into consideration but can use more abstract mental states, relating, for example, to groups of neurons. This is a way to exploit within the modeling area results from the large and more and more growing amount of neurological literature. This can be considered as lifting neurological knowledge to a mental (cognitive/affective) level; see also Chap. 2.

Within Social Neuroscience it is often suggested that innate factors may play a role, but also that a mirror neuron system can only function after a learning process has taken place (e.g., Catmur et al. 2007; Iacoboni 2008a, b; Keysers and Perrett 2004): the strength of a mirror neuron system may change over time within one person. In this chapter, in Sect. 9.5 and further an adaptive temporal-causal network model was presented addressing this aspect of adaptation over time, again based on knowledge from Social Neuroscience. The learning mechanism used is based on Hebbian learning, as also suggested by Keysers and Perrett (2004). It is shown how under normal conditions by learning the empathic responses become better over time, provided that a certain amount of social encounters occur. The model also shows how imitation training (e.g., Field et al. 2001; Ingersoll et al. 2007) can strengthen the empathic responses. Moreover, it shows that when enhanced sensory processing sensitivity (Baker et al. 2008) occurs (e.g., as an innate factor), the natural learning process is obstructed by avoidance behaviour to downregulate the dysproportional arousal (Gross 1998).

In Oztop and Arbib (2002) a computational model for a mirror neuron system for grasp actions is presented; learning is also incorporated, but in a biologically implausible manner, as also remarked in Keysers and Perrett (2004). In contrast, the presented model is based on a biologically plausible Hebbian learning model, as also suggested by Keysers and Perrett (2004).

References

- A.E.Z. Baker, A.E. Lane, M.T. Angley, R.L. Young, The relationship between sensory processing patterns and behavioural responsiveness in autistic disorder: a pilot study. *J. Autism Dev. Disord.* **38**(867–875), 2008 (2008)
- J. Bates, A.B. Loyall, W.S. Reilly, An architecture for action, emotion, and social behavior, in *Proceedings of the 4th European Workshop on Modelling Autonomous Agents in a Multi-Agent World, MAAMAW'92, Selected Papers*, ed. by C. Castelfranchi, E. Werner, vol. 830, Lecture Notes in Computer Science (Springer, 1994), pp. 55–68
- G. Bi, M. Poo, Synaptic modification by correlated activity: Hebb's postulate revisited. *Annu. Rev. Neurosci.* **24**, 139–166 (2001)
- T. Bickmore, R. Fernando, L. Ring, D. Schulman, Empathic touch by relational agents. *IEEE Trans. Affect. Comput.* **1**(60–71), 2010 (2010)
- T.W. Bickmore, R.W. Picard, Towards caring machines, in *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI)*, ed. by E. Dykstra-Erickson, M. Tscheligi (2004), pp. 1489–1492
- K.A. Bonvicini, M.J. Perlin, C.L. Bylund, G.A. Carroll, R.A. Rouse, M.G. Goldstein, Impact of communication training on physician expression of empathy in patient encounters. *Patient Educ. Counsel.* **75**, 3–10 (2009)
- T. Bosse, Z.A. Memon, J. Treur, A cognitive and neural model for adaptive emotion reading by mirroring preparation states and Hebbian learning. *Cogn. Syst. Res.* **12**(2012), 39–58 (2012)
- H. Boukricha, I. Wachsmuth, Empathy-based emotional alignment for a virtual human: a three step approach. *Künstliche Intelligenz* **25**, 195–204 (2011)
- M. Brass, S. Spengler, The inhibition of imitative behaviour and attribution of mental states, in *Social Cognition: Development, Neuroscience, and Autism*, ed. by T. Striano, V. Reid (Wiley-Blackwell, 2009), pp 52–66

- J.T. Cacioppo, G.G. Berntson, *Social Neuroscience* (Psychology Press, 2005)
- J.T. Cacioppo, P.S. Visser, C.L. Pickett, *Social neuroscience: people thinking about thinking people* (MIT Press, Cambridge, MA, 2006)
- C. Catmur, V. Walsh, C. Heyes, Sensorimotor learning configures the human mirror system. *Curr. Biol.* **17**, 1527–1531 (2007)
- L. Crane, L. Goddard, L. Pring, Sensory processing in adults with autism spectrum disorders. *Autism* **13**, 215–228 (2009)
- B. Corden, R. Chilvers, D. Skuse, Avoidance of emotionally arousing stimuli predicts social-perceptual impairment in Asperger syndrome. *Neuropsychologia* **46**(137–147), 2008 (2008)
- A.R. Damasio, *The Feeling of What Happens. Body and Emotion in the Making of Consciousness* (Harcourt Brace, New York, 1999)
- A.R. Damasio, *Looking for Spinoza* (Harcourt, 2003)
- M. Dapretto, M.S. Davies, J.H. Pfeifer, A.A. Scott, M. Sigman, S.Y. Bookheimer, M. Iacoboni, Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorder. *Nat. Neurosci.* **9**(28–30), 2006 (2006)
- A.F. de Hamilton, R.M. Brindley, U. Frith, Imitation and action understanding in autistic spectrum disorders: How valid is the hypothesis of a deficit in the mirror neuron system? *Neuropsychologia* **45**, 1859–1868 (2007)
- F. De Vignemont, T. Singer, The empathic brain: how, when and why? *Trends Cogn. Sci.* **10**(437–443), 2006 (2006)
- J. Decety, J.T. Cacioppo (eds.), *Handbook of Social Neuroscience* (Oxford University Press, 2010)
- T. Field, C. Sanders, J. Nadel, Children with autism display more social behaviors after repeated imitation sessions. *Autism* **5**, 317–323 (2001)
- I. Fried, R. Mukamel, G. Kreiman, Internally generated preactivation of single neurons in human medial frontal cortex predicts volition. *Neuron* **69**, 548–562 (2011)
- U. Frith, *Autism, Explaining the Enigma* (Blackwell, 2003)
- B. Gepner, F. Féron, Autism: a world changing too fast for a mis-wired brain? *Neurosci. Biobehav. Rev.* **33**, 1227–1242 (2009)
- W. Gerstner, W.M. Kistler, Mathematical formulations of Hebbian learning. *Biol. Cybern.* **87**, 404–415 (2002)
- P.R. Goldin, K. McRae, W. Ramel, J.J. Gross, The neural bases of emotion regulation: reappraisal and suppression of negative emotion. *Biol. Psychiat.* **63**, 577–586 (2008)
- J. Gratch, S. Marsella, P. Petta, Modeling the antecedents and consequences of emotion. *Cogn. Syst. Res.* **10**(1–5), 2009 (2009)
- J. Grèzes, B. de Gelder, Social perception: understanding other people's intentions and emotions through their actions, in *Social Cognition: Development, Neuroscience, and Autism*, ed. by T. Striano, V. Reid (Wiley-Blackwell, 2009), pp. 67–78
- J. Grèzes, B. Wicker, S. Berthoz, B. de Gelder, A failure to grasp the affective meaning of actions in autism spectrum disorder subjects. *Neuropsychologica* **47**(1816–1825), 2009 (2009)
- J.J. Gross, Antecedent- and response-focused emotion regulation: divergent consequences for experience, expression, and physiology. *J. Personality Social. Psych.* **74**, 224–237 (1998)
- J.J. Gross, Emotion regulation in adulthood: timing is everything. *Curr. Directions Psychol. Sci.* **10**, 214–219 (2001)
- J.J. Gross (ed.), *Handbook of Emotion Regulation* (Guilford Press, New York, 2007)
- E. Harmon-Jones, P. Winkielman (eds.), *Social neuroscience: Integrating biological and psychological explanations of social behavior* (Guilford, New York, 2007)
- D. Hebb, *The Organisation of Behavior* (Wiley, 1949)
- M. Hendriks, J. Treur, Modeling super mirroring functionality in action execution, imagination, mirroring, and imitation, in *Proceedings of the Second International Conference on Computational Collective Intelligence, ICCCI'10, Part I*, vol. 6421, Lecture Notes in Artificial Intelligence, ed. by J.-S. Pan, et al. (Springer, 2010), pp. 330–342
- M. Hojat, *Empathy in patient care: antecedents, development, measurement, and outcomes* (Springer, New York, 2007)

- M. Hojat, Ten approaches for enhancing empathy in health and human services cultures. *J. Health Human Services Adm.* **31**, 412–450 (2009)
- C. Hutt, S.J. Hutt, D. Lee, C. Ousted, Arousal and childhood autism. *Nature* **204**, 908–909 (1964)
- M. Iacoboni, *Mirroring People: The New Science of How We Connect with Others* (Farrar, Straus & Giroux, 2008)
- M. Iacoboni, Mesial frontal cortex and super mirror neurons. *Behav. Brain Sci.* **31**, 30 (2008b)
- M. Iacoboni, M. Dapretto, The mirror neuron system and the consequences of its dysfunction. *Nat. Rev. Neurosci.* **7**, 942–951 (2006)
- B. Ingersoll, E. Lewis, E. Kroman, Teaching the imitation and spontaneous use of descriptive gestures in young children with autism using a naturalistic behavioral intervention. *J. Autism. Dev. Disord.* **37**, 1446–1456 (2007)
- C. Keysers, V. Gazzola, Social neuroscience: mirror neurons recorded in humans. *Curr. Biol.* **20** (2010), 253–254 (2010)
- C. Keysers, D.I. Perrett, Demystifying social cognition: a Hebbian perspective. *Trends Cogn. Sci.* **8**, 501–507 (2004)
- J.C. Kirchner, A. Hatri, H.R. Heekeren, I. Dziobek, Autistic symptomatology, face processing abilities, and eye fixation patterns. *J. Autism Dev. Disord.* **41**, 158–167 (2011)
- J. Klein, Y. Moon, R. Picard, This computer responds to user frustration: theory, design, results, and implications. *Interact. Comput.* **14**(119–140), 2002 (2002)
- A. Kylläinen, J.K. Hietanen, Skin conductance responses to another person's gaze in children with autism. *J. Autism Dev. Disord.* **36**, 517–525 (2006)
- A.E. Lane, R.L. Young, A.E.Z. Baker, M.T. Angley, Sensory processing subtypes in autism: association with adaptive behavior. *J. Autism Dev. Disord.* **40**, 112–122 (2010)
- I. Leite, A. Pereira, G. Castellano, S. Mascarenhas, C. Martinho, A. Paiva, Modelling empathy in social robotic companions, in *UMAP 2011 Workshops*, LNCS 7138, ed. by L. Ardisson and T. Kuflik (2012), pp. 135–147
- S. McQuiggan, J. Robison, R. Phillips, J. Lester, Modeling parallel and reactive empathy in virtual agents: an inductive approach, in *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multi-Agent Systems*, pp. 167–174 (2008)
- R. Mukamel, A.D. Ekstrom, J. Kaplan, M. Iacoboni, I. Fried, Single-neuron responses in humans during execution and observation of actions. *Curr. Biol.* **20**(750–756), 2010 (2010)
- D. Neumann, M.L. Spezio, J. Piven, R. Adolphs, Looking you in the mouth: abnormal gaze in autism resulting from impaired top-down modulation of visual attention. *Soc. Cognit. Affect. Neurosci.* **1**(194–202), 2006 (2006)
- M. Ochs, C. Pelachaud, D. Sadek, An empathic virtual dialog agent to improve human-machine interaction, in *Proceedings of 7th International Conference on Autonomous Agents and Multiagent Systems, AAMAS'08*, ed. by Padgham, Parkes, Müller Parsonspp, pp. 89–96 (2008)
- E. Oztop, M.A. Arbib, Schema design and implementation of the grasp-related mirror neuron system. *Biol. Cybern.* **87**, 116–140 (2002)
- A. Paiva, Empathy in social agents. *Int. J. Virtual Reality* **10**, 65–68 (2011)
- L. Pessoa, On the relationship between emotion and cognition. *Nat. Rev. Neurosci.* **9**(2008), 148–158 (2008)
- E.A. Phelps, Emotion and cognition: insights from studies of the human Amygdala. *Annu. Rev. Psychol.* **2006**(57), 27–53 (2006)
- J.A. Pineda, (ed.), *Mirror Neuron Systems: the Role of Mirroring Processes in Social Cognition* (Humana Press Inc., 2009)
- J. Richer, S. Coates, (ed.), *Autism, The Search for Coherence* (Jessica Kingsley Publishers, London, 2001)
- G. Rizzolatti, C. Sinigaglia, *Mirrors in the Brain: How Our Minds Share Actions and Emotions* (Oxford University Press, 2008)
- S.H. Rodrigues, S. Mascarenhas, J. Dias, A. Paiva, I can feel it too!: Emergent empathic reactions between synthetic characters, in *3rd International Conference on Affective Computing and Intelligent Interaction and Workshops (AACII)* (IEEE Press, 2009)

- T. Singer, S. Leiberg, Sharing the emotions of others: the neural bases of empathy, in *The Cognitive Neurosciences*, 4th ed., ed. by M.S. Gazzaniga (MIT Press, 2009), pp. 973–986
- A. Smith, The empathy imbalance hypothesis of autism: a theoretical approach to cognitive and emotional empathy in autistic development. *Psychol. Record* **59**, 489–510 (2009)
- M.L. Spezio, R. Adolphs, R.S.E. Hurley, J. Piven, Analysis of face gaze in autism using ‘Bubbles’. *Neuropsychologia* **45**(144–151), 2007 (2007)
- A.L. Suchman, K. Markakis, H.B. Beckman, R. Frankel, A model of empathic communication in the medical interview. *J. Amer. Med. Assoc. (JAMA)* **277**, 678–682 (1997)
- E.A. Tinbergen, N. Tinbergen, Early childhood autism: an ethological approach. *Advances in Ethology*, vol. 10. *Journal of Comparative Ethology Supplement* (Paul Perry, Berlin, 1972)
- J. Treur, A cognitive agent model displaying and regulating different social response patterns, in *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence, IJCAI’11*, ed. by T. Walsh (AAAI Press, 2011a), pp. 1743–1749
- J. Treur, A computational agent model for Hebbian learning of social interaction, in *Proceedings of the 18th International Conference on Neural Information Processing, ICONIP’11, Part I*, vol. 7062, *Lecture Notes in Artificial Intelligence*, ed. by B.-L. Lu, L. Zhang, J. Kwok (Springer, Berlin, 2011b), pp. 9–19
- J. Treur, Displaying and regulating different social response patterns: a computational agent model. *Cognit. Comput. J.* **6**(2014), 182–199 (2014)
- J.A. Tulskey, R.M. Arnold, S.C. Alexander, M.K. Olsen, A.S. Jeffreys, K.L. Rodriguez, C.S. Skinner, D. Farrell, A.P. Abernethy, K.I. Pollak, Enhancing communication between oncologists and patients with a computer-based training program: a randomized trial. *Ann. Intern. Med.* **155**, 593–602 (2011)
- Y. van der Laan, J. Treur, An agent model for computational analysis of mirroring dysfunctioning in autism spectrum disorders, in *Proceedings of the 24th International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems, IEA/AIE’11*, vol. 6703, Part I. *Lecture Notes in Artificial Intelligence*, ed. by K.G. Mehrotra et al. (Springer, 2011), pp. 306–316
- L. van der Lubbe, J. Treur, W. van Vught, Modeling internalizing and externalizing behaviour in autism spectrum disorders, in *Proceedings of the 8th International Conference in Computational Collective Intelligence, ICCCI’16*, *Lecture Notes in AI*, ed. by N.T. Nguyen, et al. (Springer, 2016)
- C. von Hofsten, G. Gredebäck, The role of looking in social cognition: perspectives from development and autism, in *Social Cognition: Development, Neuroscience, and Autism*, ed. by T. Striano, V. Reid (Wiley-Blackwell, 2009), pp. 237–253
- J.H. Williams, A. Whiten, T. Suddendorf, D.I. Perrett, Imitation, mirror neurons and autism. *Neurosci. Biobehav. Rev.* **25**, 287–295 (2001)
- H. Yang, Z. Pan, M. Zhang, C. Ju, Modeling emotional action for social characters. *Knowl. Eng. Rev.* **23**(321–337), 2008 (2008)
- C. Zimmermann, L. Del Piccolo, J. Benzing et al., Coding patient emotional cues and concerns in medical consultations: the Verona coding definitions of emotional sequences (VR-CoDES). *Patient Educ. Couns.* **82**, 141–148 (2011)

Chapter 10

Are You with Me? Am I with You?

Joint Decision Making Processes Involving Emotion-Related Valuing and Mutual Empathic Understanding

Abstract In this chapter, based on a Network-Oriented Modeling approach a temporal-causal network model for joint decision making is presented addressing the role of mutually acknowledge empathic understanding in the decision making. The model is based on principles from recent neurological theories on mirror neurons, internal simulation, and emotion-related valuing. Emotion-related valuing of decision options and mutual contagion of intentions and emotions between persons are used as a basis for mutual empathic understanding and convergence of decisions and their associated emotions.

10.1 Introduction

An important aspect in group functioning is the ability for joint decision making. In recent years developments in neuroscience have clarified some of the mechanisms underlying such processes (e.g., Cacioppo and Berntson 2005; Decety and Cacioppo 2010; Harmon-Jones and Winkielman 2007). Two interrelated core concepts in this discipline are mirror neurons and internal simulation. Mirror neurons are neurons that not only have the function to prepare for a certain action or body change, but are also activated upon observing somebody else who is performing or tending to perform this action or body change (e.g., Iacoboni 2008a; Pineda 2009; Rizzolatti and Sinigaglia 2008). Internal simulation is mental processing that copies processes that may take place externally, for example, in another individual (e.g., Damasio 1994, 2003; Gallese and Goldman 1998; Goldman 2006; Hesslow 1994, 2002, 2012). On the one hand, mirror neurons and internal simulation have been put forward as a basic mechanism for imitation and contagion of actions and emotions; on the other hand, they have been related to empathy; e.g., (Iacoboni 2008a). In this way mirror neurons and internal simulation provide a basis both to mutually tune individual intentions and emotions and to develop mutual empathic understanding between persons (e.g., Gallese and Goldman 1998; Goldman 2006; Preston and de Waal 2002; Singer and Leiberg 2009). Usually

these two aspects are addressed separately, but in joint decision making processes they both play their roles in order to achieve solidly grounded joint decisions.

Empathic understanding can concern both cognitive (e.g., knowing or believing) and affective (e.g., feeling) aspects (see also Chap. 7, Sect. 7.3.2). Affective and cognitive understanding are often related to each other, as any cognitive state triggers an associated emotional response which is the basis of the related feeling (e.g., Damasio 1994, 1999, 2003, 2010). Usually in an individual decision making process, before a decision option is chosen an internal simulation takes place to predict the expected effects of the option (e.g., Becker and Fuchs 1985; Damasio 1994, 1999, 2003, 2010; Moore and Haggard 2008). Based on these predicted effects a valuation of the option takes place, which may involve or even be mainly based on the affective state associated to this effect (e.g., Bechara et al. 2003; Damasio 1994, 1996, 1999; Morrison and Salzman 2010; Murray 2007; Jenison et al. 2011; Ho et al. 2012; Ruff and Fehr 2014; Janak and Tye 2015; Chang et al. 2015). To achieve a solid joint decision, a shared feeling and valuation for the chosen option are important, and also mutual recognition of this sharedness. When this is achieved, a common decision has a strong shared emotional grounding as the group members do not only intend to follow that option, but they also share a good feeling about it, and they have (mutually acknowledged) empathic understanding of how other persons feel about the options. The latter may be important as well for acceptance of non-joint decisions.

The process of joint decision making was modeled using a Network-Oriented Modeling approach based on temporal-causal networks. The obtained temporal-causal network model can be used as a basis for the design of human-like virtual agents for simulation-based training or in gaming, or for virtual stories. For the first type of application the idea is to develop a number of virtual agents cooperating with a human trainee as a team in an decision making task. For the second type of application the idea is to design a system for virtual stories in which, for example, persons play a role which can be based on the presented model.

In this chapter, first in Sect. 10.2 some core concepts used are briefly reviewed. Next, in Sect. 10.3 the temporal-causal network model is presented. In Sect. 10.4 some of the explored simulation scenarios are discussed. Finally, Sect. 10.5 is a discussion.

10.2 Mirroring, Internal Simulation and Emotion-Related Valuing

Two concepts used here as a basis are mirror neurons and internal simulation; in combination they provide an individual's mental function of mirroring mental processes of another individual (see also Chap. 7). Mirror neurons are not only firing when a subject is preparing for an action, but also when somebody else is performing or preparing for this action and the subject just observes that. They have first been found in monkeys (e.g., Gallese et al. 1996; Rizzolatti et al. 1996), and after that it has been assumed that similar types of neurons also occur in humans,

with empirical support, for example, in (Iacoboni et al. 2005) based on fMRI, and (Fried et al. 2011; Mukamel et al. 2010) based on single cell experiments with epilepsy patients; see also (Keysers and Gazzola 2010). For reviews, see (Molenberghs et al. 2012; Kilner and Lemon 2013). The effect of activation of mirror neurons is context-dependent. A specific type of neurons has been suggested to be able to indicate such a context. They are assumed to indicate self-other distinction and exert control by allowing or suppressing action execution; e.g., Brass and Spengler 2009; Iacoboni 2008b), and (Iacoboni 2008a, pp. 196–203).

Activation states of mirror neurons play an important role in *mirroring* mental processes of other persons, which is based on *mirror neuron activation* in combination with *internal simulation*. A body loop is the following causal chain for generation of felt emotions; see (James 1884; Damasio 2010, pp. 114–116):

sensory representation → preparation for bodily changes → expressed bodily changes →
sensory representation of (sensed) bodily changes

In an internally simulated form an *as-if body loops* is bypassing actually expressed bodily changes; e.g., (Damasio 1994, pp. 155–158, 1999, pp. 79–80):

sensory representation → preparation for bodily changes →
sensory representation of (simulated) bodily changes

So, an as-if body loop describes an *internal simulation* of the bodily processes, without actually affecting the body, comparable to simulation in order to perform, for example, prediction, mindreading or imagination; e.g., (Becker and Fuchs 1985; Gallese and Goldman 1998; Goldman 2006; Hesslow 2002; Moore and Haggard 2008). The feelings generated in this way play an important role in valuing predicted or imagined effects of actions, in relation to amygdala activations; see, e.g., (Morrison and Salzman 2010; Murray 2007). The emotional response and feeling mutually affect each other in a bidirectional manner: an as-if body loop usually has a cyclic form by a connection back; see, for example, (Damasio 2003, pp. 91–92, 2010, pp. 119–122):

sensory representation of (simulated) bodily changes → preparation for bodily changes

As mirror neurons make that some specific sensory input (an observed action of another person) directly links to related preparation states, they combine well with internal simulation in the form of as-if body loops; see also (Damasio 2010, pp. 102–104). In this way states of other persons lead to activation of some of a person's corresponding own states that at the same time play a role in the person's own feelings and decisions for actions. This provides an effective mechanism for how observed actions and feelings and own actions and feelings are tuned to each other. Thus a mechanism is obtained which explains how in a social context persons fundamentally affect each other's individual decisions and states, including feelings. Moreover, it is also the basis for empathic understanding of other persons' preferences and feelings. Both the tuning and convergence of action tendencies and the mutual empathic understanding (even when finally no common option is decided for) play a crucial role in joint decision making processes.

10.3 The Temporal-Causal Network Model

The issues and perspectives briefly reviewed in the introduction and Sect. 10.2 have been used as a basis for the neurologically inspired temporal-causal network model presented below (for a conceptual representation, see Fig. 10.1). First, the following criteria for a solid joint decision have been taken as a point of departure:

Solid joint decision

The outcome of a joint decision process is a *solid joint decision* if the following criteria are fulfilled:

- a *common action* option choice
- a *shared positive feeling* and *valuation* for the effect of this action option
- mutually *acknowledged empathic understanding* for both the action and feeling

The idea is that decisions that do not satisfy these criteria have a higher chance of being revised as soon as a slight opportunity occurs for that. A second point of departure is emotion-related valuing:

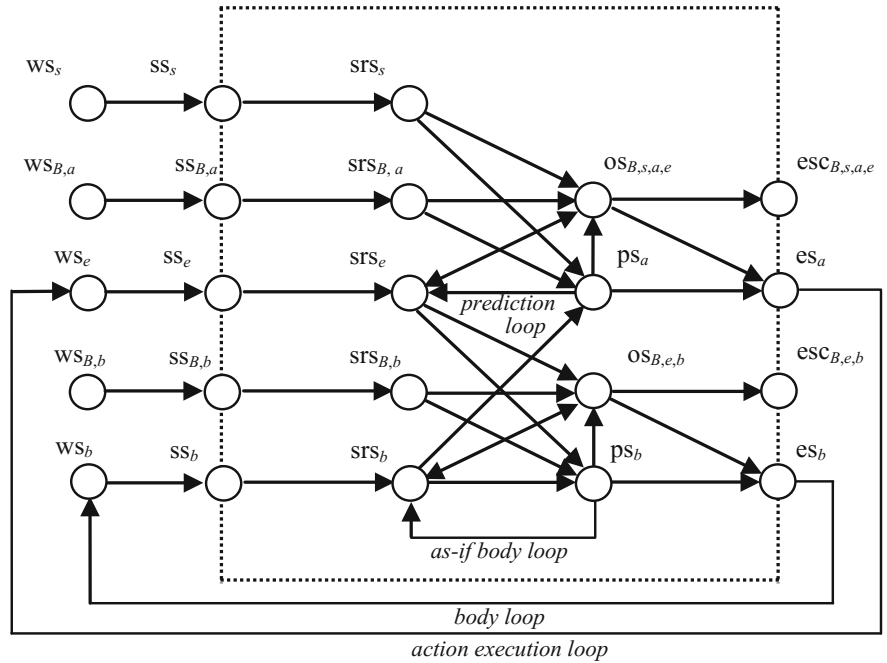


Fig. 10.1 Graphical conceptual representation of the temporal-causal network model

Emotion-related valuing of options

The choice for a specific action option is based on *emotion-related valuing* of the *predicted effects* of the action options; see also Chap. 6, Sect. 6.2.

As a third point of departure it is assumed that mirroring takes place both for the action options and for the emotions:

Mirroring both of actions and emotions

Both the tendency to go for an action option and the associated emotion are transferred between persons via *mirroring processes* using a *mirror neuron* function and *internal simulation*

These mirroring processes at the same time induce a gradual process of mutually *tuning* the considered actions and their emotion-related valuations, and the development of mutual *empathic understanding*. So, finally, a fourth notion that has been used as a point of departure is mutually acknowledged empathic understanding:

Mutually acknowledged empathic understanding

Mutually acknowledged empathic understanding is based on the following criteria:

- (a) Showing the same state as the other person (nonverbal part of the empathic response)
- (b) Telling that the other person has this state (verbal part of the empathic response)

Assuming true, faithful nonverbal and verbal expression, these criteria are in line with the criteria of empathy for affective states formulated in (Singer and Leiberg 2009).

10.3.1 Conceptual Representation of the Temporal-Causal Network Model

In the model s denotes a *stimulus*, a an *option* for an *action* to be decided about, and e a world state which is an *effect* of the action. The effect state e is *valued* by associating a *feeling* state b to it, which is considered to be positive for the person (e.g., in accordance with a goal). The states used in the model are summarized in Table 10.1.

The temporal-causal network model uses ownership states for actions a and their effects e , both for self and other persons, specified by $os_{B,s,a,e}$ with B another person or self, respectively (see Fig. 10.1). Similarly, ownership states are used for emotions indicated by body state b , both for self and other persons, specified by $os_{B,e,b}$ with B another person or self. As an example, the four arrows to $os_{B,s,a,e}$ in Fig. 10.1 show that an ownership state $os_{B,s,a,e}$ is affected by the preparation state ps_a for the action a , the sensory representation srs_b of the emotion-related value b for the predicted effect

Table 10.1 States used

Notation	Description
ws_W	World state W : for an action a of person B , a feeling b of person B , a stimulus s , effect e , or an emotion indicated by body state b
ss_W	Sensor state for W
srs_W	Sensory representation of W
ps_X	Preparation state for X : action a or expressing emotion by body state b
es_X	Execution state for X : action a or expressing emotion by body state b
$os_{B,s,a,e}$	Ownership state for B of action a with effect e and stimulus s
$os_{B,e,b}$	Ownership state for B of emotion indicated by body state b and effect e
$esc_{B,s,a,e}$	Communication to B of ownership for B of action a with effect e and stimulus s
$esc_{B,e,b}$	Communication to B of ownership for B of emotion indicated by b and effect e

e , the sensory representation srs_s of the stimulus s , and the sensory representation srs_B of the person B . Note that s , a , e , b , and B are parameters for stimuli, actions, effects, body states, and persons. In a given model multiple instances of each of them can occur.

Prediction of effects of prepared actions is modeled using the connection from the preparation ps_a of the action a to the sensory representation srs_e of the effect e . Suppression of the sensory representation of a predicted effect (according to, e.g., Blakemore et al. 1999, 2000; Moore and Haggard 2008) is modeled by the (inhibiting) connection from the ownership state $os_{B,s,a,e}$ to sensory representation srs_e . The control exerted by the ownership state for action a is modeled by the connection from $os_{B,s,a,e}$ to es_a . Communicating ownership for an action (a way of expressing recognition of the other person's states, as a verbal part of showing empathic understanding) is modeled by the connection from the ownership state $os_{B,s,a,e}$ to the communication execution state $esc_{B,s,a,e}$. Similarly, communicating ownership for an emotion for effect e indicated by b is modeled by the connection from the ownership state $os_{B,e,b}$ to the communication execution state $esc_{B,e,b}$. Connections between states (the arrows in Fig. 10.1) have weights, as indicated in Table 10.2.

In this table the column LP refers to the (temporally) Local Properties LP1 to LP9 presented as a numerical representation in difference equation format below. A connection weight usually has a value between -1 and 1 and may depend on the specific instance for person B , stimulus s , action a and/or effect state b involved. Note that in general weights are assumed non-negative, except for inhibiting connections, such as $\omega_{2,2,e}$ which models suppression of the sensory representation of effect e , and $\omega_{2,2,b}$ which models suppression of the sensory representation of body state b .

Table 10.2 Overview of the connections and their weights

From states	To state	Weight name	Connection name	LP
ss_X	srs_X	$\omega_{1,X}$	Representing	LP1
ps_a	srs_e	$\omega_{2,1,e}$	Predicting e	LP2
$os_{B,s,a,e}$		$\omega_{2,2,e}$	Controlling	
ss_e		$\omega_{2,3,e}$	Representing e	
ps_b		$\omega_{2,1,b}$	Predicting b	
$os_{B,e,b}$	srs_b	$\omega_{2,2,b}$	Suppressing	
ss_b		$\omega_{2,3,b}$	Representing b	
srs_s				
srs_b	ps_a	$\omega_{3,1,a}$	Responding to s	LP3
$srs_{B,a}$		$\omega_{3,2,a}$	Amplifying a	
srs_e		$\omega_{3,3,a}$	Responding to a of B	
srs_e	ps_b	$\omega_{3,1,b}$	Responding to e	
srs_b		$\omega_{3,2,b}$	Amplifying b	
$srs_{B,b}$		$\omega_{3,3,b}$	Responding to b of B	
$srs_{B,a}$	$os_{B,s,a,e}$	$\omega_{4,1,a}$	Owning a for B and a	LP4
srs_s		$\omega_{4,2,a}$	Owning a for s	
ps_a		$\omega_{4,3,a}$	Owning a for a	
srs_e		$\omega_{4,4,a}$	Owning a for e	
$srs_{B,b}$	$os_{B,e,b}$	$\omega_{4,1,b}$	Owning b for B and b	
srs_e		$\omega_{4,2,b}$	Owning b for e	
ps_b		$\omega_{4,3,b}$	Owning b for b preparation	
srs_b		$\omega_{4,4,b}$	Owning b for b representation	
$os_{B,s,a,e}$	es_a	$\omega_{5,1,a}$	Controlling a	LP5
ps_a		$\omega_{5,2,a}$	Executing a	
$os_{B,e,b}$	es_b	$\omega_{5,1,b}$	Controlling b	
ps_b		$\omega_{5,2,b}$	Executing b	
es_a	ws_e	$\omega_{6,e}$	Effectuating e	LP6
es_b	ws_b	$\omega_{6,b}$	Effectuating b	
ws_X	ss_X	$\omega_{7,X}$	Sensing X	LP7
$os_{B,s,a,e}$	$esc_{B,s,a,e}$	$\omega_{8,a}$	Executing communication of a to B	LP8
$os_{B,e,b}$	$esc_{B,e,b}$	$\omega_{8,b}$	Executing communication of b to B	

10.3.2 Numerical Representation of the Temporal-Causal Network Model

Below, the dynamics following the connections between the states in Fig. 10.1 are described in more detail. This is done for each state by a dynamic property specifying how the activation value for this state is updated based on the activation values of the states connected to it (the incoming arrows in Fig. 10.1). Note that in these property specifications s , a , e , b , and B are parameters for stimuli, actions, effects, body states, and persons, respectively; multiple instances for each of them

can be used in a given model. Parameter η indicates the speed by which an activation level is updated based on received input from other states.

In the example simulations, in LP1, LP6, and LP7, the combination function $c(V)$ is the identity function

$$c(V) = \mathbf{id}(V) = V$$

For the other states the combination function $c(V_1, \dots, V_k)$ is the advanced logistic sum function:

$$\begin{aligned} c(V_1, \dots, V_k) &= \mathbf{alogistic}_{\sigma, \tau}(V_1, \dots, V_k) \\ &= \left(\frac{1}{1 + e^{-\sigma(V_1 + \dots + V_k - \tau)}} - \frac{1}{1 + e^{\sigma\tau}} \right) (1 + e^{-\sigma\tau}) \end{aligned}$$

Other types of combination functions might be used as well. For example values for τ and σ , see in Sect. 10.4.

The first property LP1 describes how sensory representations are generated for any state X , indicating a stimulus s , an action a of a person B , or a feeling b of a person B .

LP1 Sensory representation of X based on a sensor state for X

$$\text{srs}_X(t + \Delta t) = \text{srs}_X(t) + \eta_{\text{srs}_X} [\text{c}_{\text{srs}_X}(\omega_{1,X}\text{ss}_X(t)) - \text{srs}_X(t)]\Delta t$$

The sensory representation of an effect state e is not only affected by a corresponding sensor state for e (affected by the world state), but also by two action-related states:

- via the *predictive loop* by a preparation state, as a way of *internal simulation* to predict the effect e of a prepared action a
- by an inhibiting connection from the self-ownership state, to *suppress* the sensory representation of the *effect* e of the action a , once it is going to be initiated; e.g., Blakemore et al. (1999, 2000)

This is expressed in dynamic property LP2. Note that for this suppressing effect the connection weight $\omega_{2,2,e}$ from ownership state for action a to sensory representation for effect e is chosen negative, for example $\omega_{2,2,e} = -0.2$. Dynamic property LP2b specifies a similar temporal relationship for update of the sensory representation of a body state, and thus models *internal simulation* by an *as-if body loop*.

LP2e Sensory representation for an effect state e

$$\text{srs}_e(t + \Delta t) = \text{srs}_e(t) + \eta_{\text{srs}_e} [\text{c}_{\text{srs}_e}(\omega_{2,1,e}\text{ps}_a(t), \omega_{2,2,e}\text{os}_{B,s,a,e}(t), \omega_{2,3,e}\text{ss}_e(t)) - \text{srs}_e(t)]\Delta t$$

LP2b Sensory representation for a body state b

$$\text{srs}_b(t + \Delta t) = \text{srs}_b(t) + \eta_{\text{srs}_b} [\text{c}_{\text{srs}_b}(\omega_{2,1,b}\text{ps}_b(t), \omega_{2,2,ob}\text{os}_{B,e,b}(t), \omega_{2,3,b}\text{ss}_b(t)) - \text{srs}_b(t)]\Delta t$$

Preparation for action a is affected by

- the sensory representation of stimulus s
- the body state b associated to the predicted effect e of the action,
- observation of the action (tendency) in another person

The first bullet is an external trigger for the action. The second bullet models the impact of the result b of the *emotion-related valuing* of the action effect e . The third bullet models the *mirroring* effect for the action as observed as a tendency in another person. Similarly for the preparation for a body state b ; here the sensory representation of the effect e serves as a trigger, and the emotion state of another person is mirrored.

LP3a Preparing for an action a

$$ps_a(t + \Delta t) = ps_a(t) + \eta_{ps_a} [c_{ps_a}(\omega_{3,1,a}sr_s(t), \omega_{3,2,a}sr_b(t), \omega_{3,3,B,a}sr_{B,a}(t)) - ps_a(t)]\Delta t$$

LP3b Preparing for a body state b

$$ps_b(t + \Delta t) = ps_b(t) + \eta_{ps_b} [c_{ps_b}(\omega_{3,1,b}sr_e(t), \omega_{3,2,b}sr_b(t), \omega_{3,3,B,b}sr_{B,b}(t)) - ps_b(t)]\Delta t$$

Ownership states for an action a or body state b are generated by LP4a and LP4b. They keep track of the person's context with respect to the action or body state. This context concerns both the person self and the other persons and their extent of ownership of the action or body change; in this sense it is a basis for attribution to a person, and includes self-other distinction. Moreover, a self-ownership is used to control execution of prepared actions or body states, like super mirror neurons are assumed to do. For example, in case the person B is *self*, the ownership state for action a strengthens the initiative to perform a as a self-generated action: executing a prepared action depends on whether a certain activation level of the ownership state for the person *self* is available for this action. This is how control over the execution of the action (go/no-go decision) is exerted, and can, for example, be used to veto the action in a stage of preparation.

LP4a Generating an ownership state for B and a

$$os_{B,s,a,e}(t + \Delta t) = os_{B,s,a,e}(t) + \eta_{os_{B,s,a,e}} [c_{os_{B,s,a,e}}(\omega_{4,1,a}sr_{B,a}(t), \omega_{4,2,a}sr_s(t), \omega_{4,3,a}ps_a(t), \omega_{4,4,a}sr_e(t)) - os_{B,s,a,e}(t)]\Delta t$$

LP4b Generating an ownership state for B and b

$$os_{B,e,b}(t + \Delta t) = os_{B,e,b}(t) + \eta_{os_{B,e,b}} [c_{os_{B,e,b}}(\omega_{4,1,b}sr_{B,b}(t), \omega_{4,2,b}sr_e(t), \omega_{4,3,b}ps_b(t), \omega_{4,4,b}sr_b(t)) - os_{B,e,b}(t)]\Delta t$$

Note that in case that B is the person self, the first condition in LP4a and LP4b indicates how far the person has a certain willingness to come to an action or expression. For example, when no other person is present the willingness to explicitly express emotions may be less, or when the person is in a passive mood, willingness to come to an action may be low. The use of ownership states in control of execution is modeled by LP5:

LP5a Execution of action a

$$es_a(t + \Delta t) = es_a(t) + \eta_{es_a} [c_{es_a}(\omega_{5,1,a}os_{B,s,a,e}(t), \omega_{5,2,a}ps_a(t)) - es_a(t)]\Delta t$$

LP5b Execution of body change b

$$es_b(t + \Delta t) = es_b(t) + \eta_{es_b} [c_{es_b}(\omega_{5,1,b}os_{B,e,b}(t), \omega_{5,2,b}ps_b(t)) - es_b(t)]\Delta t$$

Note that these executions also function as the *nonverbal part of the empathic response*; e.g., showing a face expression with the same emotion as the other person.

Property LP6 describes in a straightforward manner how execution of action a or body change b affects the world state for effect e or body state b .

LP6e From action execution to effect state

$$ws_e(t + \Delta t) = ws_e(t) + \eta_{ws_e} [c_{ws_e}(\omega_{6,e}es_a(t)) - ws_e(t)]\Delta t$$

LP6b From body change execution to resulting body state

$$ws_b(t + \Delta t) = ws_b(t) + \eta_{ws_b} [c_{ws_b}(\omega_{6,b}es_b(t)) - ws_b(t)]\Delta t$$

The following property models how sensor states are updated. It applies to an action a of person B , a feeling b of person B , a stimulus s , effect e , or emotion indicated by body state b (covered by variable X).

LP7 Generating a sensor state for a world or body state X

$$ss_X(t + \Delta t) = ss_X(t) + \eta_{ss_X} [c_{ss_X}(\omega_{7,X}ws_X(t)) - ss_X(t)]\Delta t$$

Communication of ownership of the other person to the other person represents acknowledgement of an person that it has noticed the state of the other person: a *verbal part of the empathic response*. These communications depend on the ownership states as specified in LP8.

LP8a Communication of the other person B's intention for a and e for s

$$esc_{B,s,a,e}(t + \Delta t) = esc_{B,s,a,e}(t) + \eta_{esc_{B,s,a,e}} [c_{esc_{B,s,a,e}}(\omega_{8,a}os_{B,s,a,e}(t)) - esc_{B,s,a,e}(t)]\Delta t$$

LP8b Communication of the other person B's emotion b for e

$$esc_{B,e,b}(t + \Delta t) = esc_{B,e,b}(t) + \eta_{esc_{B,e,b}} [c_{esc_{B,e,b}}(\omega_{8,b}os_{B,e,b}(t)) - esc_{B,e,b}(t)]\Delta t$$

10.4 Simulation Results

In this section simulation results are discussed for one of the scenarios that have been explored. Note that in this section for the sake of simplicity two persons A and B are considered and for each of s , a , e , b , just one instance is used, which is the same for both persons. In the scenario all connection strengths were set at 1, except the inhibiting connections, which were set at -0.2 , and the connection to the action effect in the world which was taken 0 as the focus here is on the process of decision making prior to the actual execution of the decision. The speed factor η was set to 0.5 and $\Delta t = 0.2$. In the scenario shown in Fig. 10.2 both persons get stimulus s as input with level 1. Here time is on the horizontal axis and activation levels as indicated are on the vertical axis. The upper graph shows person A and the lower graph person B.

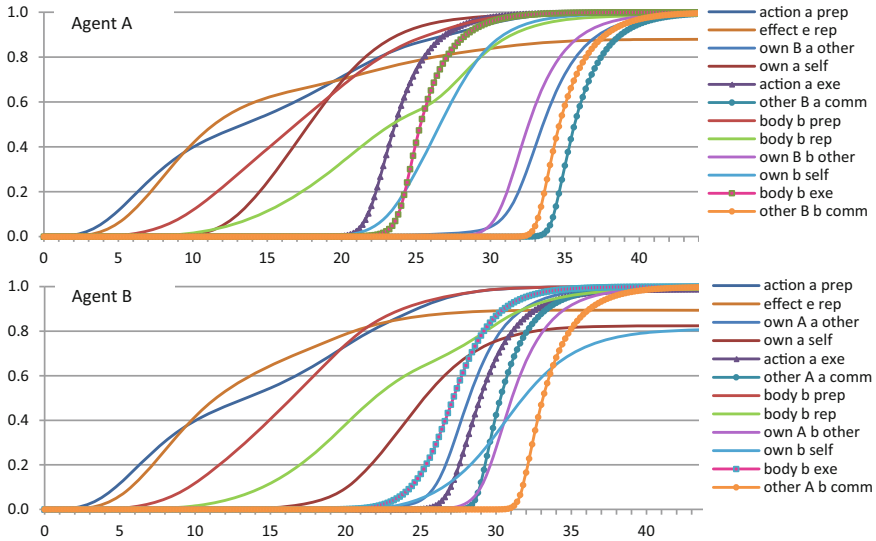


Fig. 10.2 Reaching a joint decision and mutual understanding for different self-contexts

The threshold and steepness values used are as follows: LP2e (0.2, 4), LP2b (0.7, 4), LP3a (1, 4), LP3b (0.7, 4), LP4 (3.2, 8), LP5a (1.6, 20), LP5b (1, 20), LP8 (0.6, 20).

The only difference between the two persons is that person A has level 1 for the self-context factor which indicates willingness to come to action and for person B this is 0.5. In Fig. 10.2 it is shown that triggered by the stimulus s , from time point 3 on both persons develop a preparation for action option a , which is immediately followed by activation of predicted effect e . Next, around time point 6 both persons start to develop an emotional response preparation for b on the predicted effect e , and as a consequence (by the as-if body loop) the feeling of this emotion from time point 9 on. Around time point 10 person A starts to activate the self ownership state for action option a , whereas for person B this only happens later, after time point 16, due to its lower self-context value. Due to this, person A expresses (the tendency for) action option a from time point 20 on. From time point 22 on person A expresses the emotion felt, after an ownership state for this was activated from time point 20 on. Note that at this point in time point person B does not show such reactions, due to the lower self-context for person B.

However, by B's mirroring of the two types of expression from person A (action tendency and body state for b), person B is affected in its preparation levels for both the action option and the bodily response. Due to this, person B also expresses the feeling from time point 21 and the tendency for action option a from time point 26 on. This actually creates a joint decision for action option a , accompanied by a good feeling b for it. Moreover, this also provides the nonverbal part of B's empathic response on person A's action tendency and feeling. Furthermore, person B shows a

Table 10.3 The 16 possible outcomes for each person

A acknowledges understanding of B 's intention for O	Intention acknowledgement				No intention acknowledgement											
	Feeling acknowledgement				No feeling acknowledgement				Feeling acknowledgement				No feeling acknowledgement			
	Intention		No intention		Intention		No intention		Intention		No intention		Intention		No intention	
A acknowledges understanding of B 's positive feeling for O	Feeling		No feeling		Feeling		No feeling		Feeling		No feeling		Feeling		No feeling	
	Feel	No feel	Feel	No feel	Feel	No feel	Feel	No feel	Feel	No feel	Feel	No feel	Feel	No feel	Feel	No feel
A has an intention for O	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
A has a positive feeling for O																

verbal empathic response to A for both the action and the feeling starting at time points 28 and 30, respectively. The verbal empathic response from person A comes later, at time points 32 and 33 respectively, which reflects the fact that some time was needed to get person B in the proper state (due to mirroring) to show support for action option *a* and feeling *b*. So, finally they reach a solid joint decision.

Not all joint decision making processes will achieve an outcome in the form of a solid joint decision, as in the above scenario. There are many ways in which a joint decision process can end up in a less perfect outcome. A more extensive analysis of the different possible outcomes of the joint decision making processes described by the model can be found in (Duell and Treur 2012), thereby distinguishing 16 possible outcomes for each of the persons as shown in Table 10.3.

Based on this, in principle $16^2 = 256$ possible outcomes for a joint decision making process between two persons can be distinguished. Only one of these 256 possible outcomes concerns a solid joint decision. The other 255 possible outcomes show in how many ways a joint decision process can be poor, or at least less perfect. Part of these 255 possibilities still concern a joint option choice and part of them does not. Moreover, part of the outcomes with joint option choice have a solid emotional grounding and part of them have not. Finally, some of the outcomes will lack mutual acknowledgements. For more details, see (Duell and Treur 2012). In addition, in (Duell and Treur 2012) it is analysed how the process to reach a certain outcome can differ; for example, one the persons develops a decision first and affects the other person who then follows, or both persons develop their decision simultaneously while they both mutually affect each other.

10.5 Discussion

In this chapter, a temporal-causal network model was presented which was developed based on a Network-Oriented Modeling approach, thereby using mechanisms from Social Neuroscience. The contents of this chapter are mainly based on (Treur 2011d). The model addresses the emergence of joint decisions, accompanied by shared emotions and mutually acknowledged empathic understanding. To this end it covers both cognitive and affective processes and their interaction in decision making, and social contagion. Core mechanisms adopted are mirror neurons (e.g., Iacoboni 2008a; Pineda 2009; Rizzolatti and Sinigaglia 2008), internal simulation (e.g., Damasio 1994, 2003; Gallese and Goldman 1998; Goldman 2006; Hesslow 1994, 2002, 2012), and emotion-related valuing of predicted effects of action options (e.g., Bechara et al. 2003; Damasio 1994, 1996, 2003; Morrison and Salzman 2010; Murray 2007; Jenison et al. 2011; Ho et al. 2012; Janak and Tye 2015).

It was shown how such temporal-causal network models can be used to perform simulation and analysis of the emergence of joint decisions grounded in shared emotion-related valuing, and together with mutual empathic understanding of persons.

The model uses elements from the model presented in (Treur 2011a, 2014) for the empathic understanding, but in contrast to (Treur 2011a, 2014) where the empathic understanding was limited to emotions, in the current model it is applied to both (tendencies for) actions and emotions. Furthermore, the current model uses the idea of ownership states as in the model presented in (Treur 2011b, 2012). However, in (Treur 2011b, 2012) ownership states are differentiated into prior and retrospective ownership states, which was not done in the current model. Moreover, in the current model the ownership states were used both for actions and for expressing emotions, whereas in (Treur 2011b, 2012) they were only focused on actions, and emotions were not addressed. Another difference to both (Treur 2011a) and (Treur 2011b, 2012) is the use in the current model of social contagion to affect both action tendencies and associated feelings in order to come to joint decisions accompanied by shared associated emotions.

In (Duell and Treur 2014) an extension of the model is described obtained by including a mediator model who contributes some process focusing and timing guidance. In (Duell 2014) an extension of the model is described where the persons are adaptive and a mediator guides the persons in changing their valuing of the options by some form of (Hebbian) learning in an attempt to get rid of prejudices or effects of exceptional experiences. Finally, in (Ments et al. 2015) another extension of the joint decision making model is discussed which incorporates cognitive metaphors within the persons, and how they affect the decision making process by connecting to certain states within this process; this is particularly focusing on competitive metaphors, such as war, and more cooperative metaphors.

References

- A. Bechara, H. Damasio, A.R. Damasio, Role of the Amygdala in Decision-Making. *Ann. N.Y. Acad. Sci.* **985**, 356–369 (2003)
- W. Becker, A.F. Fuchs, Prediction in the oculomotor system: smooth pursuit during transient disappearance of a visual target. *Exp. Brain Res.* **57**, 562–575 (1985)
- S.-J. Blakemore, C.D. Frith, D.M. Wolpert, Spatio-temporal prediction modulates the perception of self-produced stimuli. *J. Cog. Neurosci.* **11**, 551–559 (1999)
- S.-J. Blakemore, D.M. Wolpert, C.D. Frith, Why can't you tickle yourself? *Neuroreport* **11**, 11–16 (2000)
- M. Brass, S. Spengler, The inhibition of imitative behaviour and attribution of mental states, in *Social Cognition: Development, Neuroscience, and Autism*, ed. by T. Striano, V. Reid (Wiley-Blackwell, 2009), pp. 52–66
- J.T. Cacioppo, G.G. Berntson, *Social Neuroscience* (Psychology Press, 2005)
- S.W.C. Chang, N.A. Fagan, K. Toda, A.V. Utevsy, J.M. Pearson, M.L. Platt, Neural mechanisms of social decision-making in the primate amygdala. *PNAS* (2015)
- A.R. Damasio, (1994). *Descartes' Error: Emotion, Reason and the Human Brain* (Papermac, London, 1994)
- A.R. Damasio, The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophic. Trans. R. Soc. Biol. Sci.* **351**, 1413–1420 (1996)
- A.R. Damasio, *The Feeling of What Happens. Body and Emotion in the Making of Consciousness* (Harcourt Brace, New York, 1999)

- A.R. Damasio, *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain* (Vintage Books, London, 2003)
- A.R. Damasio, *Self Comes to Mind: Constructing the Conscious Brain* (Pantheon Books, NY, 2010)
- J. Decety, J.T. Cacioppo, (eds.), *Handbook of Social Neuroscience* (Oxford University Press, 2010)
- R. Duell, Modelling mediator intervention in joint decision making processes involving mutual empathic understanding, in *Proceedings of the 21st International Conference on Neural Information Processing, ICONIP 2014*, Lecture Notes in Computer Science, vol. 8835, ed. by C.K. Loo, K.S. Yap, K.W. Wong, A. Teoh, K. Huang (Springer International Publishing, 2014), pp. 589–596
- R. Duell, J. Treur, A computational analysis of joint decision making processes, in *Proceedings of the 4th International Conference on Social Informatics, SocInfo'12*, Lecture Notes in Computer Science, vol. 7710, ed. by K. Aberer et al. (Springer, 2012), pp. 292–308
- R. Duell, J. Treur, Modeling mediator assistance in joint decision making processes involving mutual empathic understanding, in *Proceedings of the 6th International Conference on Computational Collective Intelligence Technologies and Applications, ICCCI'14*, Lecture Notes in Artificial Intelligence, vol. 8733, ed. by D. Hwang et al. (Springer, Berlin Heidelberg, 2014), pp. 544–553
- I. Fried, R. Mukamel, G. Kreiman, Internally generated preactivation of single neurons in human medial frontal cortex predicts volition. *Neuron* **69**, 548–562 (2011)
- V. Gallese, L. Fadiga, L. Fogassi, G. Rizzolatti, Action recognition in the premotor cortex. *Brain* **119**, 593–609 (1996)
- V. Gallese, A. Goldman, Mirror neurons and the simulation theory of mindreading. *Trends Cogn. Sci.* **2**, 493–501 (1998)
- A.I. Goldman, *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading* (Oxford Univ. Press, New York, 2006)
- E. Harmon-Jones, P. Winkielman, (eds.), *Social Neuroscience: Integrating Biological and Psychological Explanations of Social Behavior* (Guilford, New York, 2007)
- G. Hesslow, Will neuroscience explain consciousness? *J. Theoret. Biol.* **171**, 29–39 (1994)
- G. Hesslow, Conscious thought as simulation of behaviour and perception. *Trends Cogn. Sci.* **6**, 242–247 (2002)
- G. Hesslow, The current status of the simulation theory of cognition. *Brain Res.* **1428**, 71–79 (2012)
- S. Ho, R.D. Gonzalez, J.L. Abelson, I. Liberzon, Neurocircuits underlying cognition–emotion interaction in a social decision making context. *NeuroImage* **63**, 843–857 (2012)
- M. Iacoboni, *Mirroring People: The New Science of How We Connect with Others* (Farrar, Straus & Giroux, New York, 2008a)
- M. Iacoboni, Mesial frontal cortex and super mirror neurons. *Beh. Brain Sci.* **31**, 30–30 (2008b)
- M. Iacoboni, I. Molnar-Szakacs, V. Gallese, G. Buccino, J.C. Mazziotta, G. Rizzolatti, Grasping the intentions of others with one's own mirror neuron system. *PLoS Biol.* **3**, e79 (2005)
- P.H. Janak, K.M. Tye, From circuits to behaviour in the amygdala. *Nature* **517**, 284–292 (2015)
- W. James, What is an emotion. *Mind* **9**, 188–205 (1884)
- R.L. Jenison, A. Rangel, H. Oya, H. Kawasaki, M.A. Howard, Value encoding in single neurons in the human amygdala during decision making. *J. Neurosci.* **31**, 331–338 (2011)
- C. Keysers, V. Gazzola, Social neuroscience: mirror neurons recorded in humans. *Curr. Biol.* **20**, 253–254 (2010)
- J.M. Kilner, R.N. Lemon, What we know currently about mirror neurons. *Curr. Biol.* **23**, R1057–R1062 (2013)
- L. Ments, D.J. van Thilakarathne, J. Treur, Modelling the role of cognitive metaphors in joint decision making, in *Proceedings of the 15th International Conference on Intelligent Agent Technology, IAT'15*, vol. 2 (IEEE Computer Society Press), pp 67–75
- P. Molenberghs, R. Cunnington, J.B. Mattingley, Brain regions with mirror properties: a meta-analysis of 125 human fMRI studies. *Neurosci. Biobehav. Rev.* **36**, 341–349 (2012)

- J. Moore, P. Haggard, Awareness of action: inference and prediction. *Consci. Cogn.* **17**, 136–144 (2008)
- S.E. Morrison, C.D. Salzman, Re-valuing the amygdala. *Curr. Opin. Neurobiol.* **20**, 221–230 (2010)
- R. Mukamel, A.D. Ekstrom, J. Kaplan, M. Iacoboni, I. Fried, Single-neuron responses in humans during execution and observation of actions. *Curr. Biol.* **20**, 750–756 (2010)
- E.A. Murray, The amygdala, reward and emotion. *Trends Cogn. Sci.* **11**, 489–497 (2007)
- J.A. Pineda, (ed.), *Mirror Neuron Systems: the Role of Mirroring Processes in Social Cognition* (Humana Press Inc., 2009)
- S.D. Preston, F.B.M. Waal de, Empathy: its ultimate and proximate bases. *Behav. Brain Sci.* **25**, 1–72 (2002)
- G. Rizzolatti, L. Fadiga, V. Gallese, L. Fogassi, Premotor cortex and the recognition of motor actions. *Cogn. Brain Res.* **3**, 131–141 (1996)
- G. Rizzolatti, C. Sinigaglia, *Mirrors in the Brain: How Our Minds Share Actions and Emotions* (Oxford University Press, 2008)
- C.C. Ruff, E. Fehr, The neurobiology of rewards and values in social decision making. *Nat. Rev. Neurosci.* **15**, 549–562 (2014)
- T. Singer, S. Leiberg, Sharing the emotions of others: the neural bases of empathy, in *The Cognitive Neurosciences*, 4th edn., ed. by M.S. Gazzaniga (MIT Press, 2009), pp. 973–986
- J. Treur, A cognitive agent model displaying and regulating different social response patterns, in *Proceedings of IJCAI'11*, ed. by T. Walsh (2011a), pp. 1735–1742
- J. Treur, A cognitive agent model incorporating prior and retrospective ownership states for actions, in *Proceedings of IJCAI'11*, ed. by T. Walsh (2011b), pp. 1743–1749
- J. Treur, Modelling joint decision making processes involving emotion-related valuing and mutual empathic understanding, in *Agents in Principle, Agents in Practice, Proceedings of the 14th International Conference on Principles and Practice of Multi-Agent Systems, PRIMA'11*. Lecture Notes in Artificial Intelligence, vol. 7047, ed. by D. Kinny, J.Y. Hsu, G. Governatori, A.K. Ghose (Springer, Berlin Heidelberg, 2011d), pp. 410–423
- J. Treur, A computational agent model incorporating prior and retrospective ownership states for actions. *Biol. Insp. Cogn. Architect. J.* **2**, 54–67 (2012)
- J. Treur, Displaying and regulating different social response patterns: a computational agent model. *Cogn. Comput. J.* **6**, 182–199 (2014)

Chapter 11

Changing Yourself, Changing the Other, or Changing Your Connection

Integrative Dynamics of States and Interactions in a Social Context

11.1 Introduction

When persons function in a social context they interact with a number of other persons they know. These persons themselves also interact with a number of persons. And so on and on. When each person is modeled by a node and for each of these interactions arcs between the nodes are drawn, this results in a social network model, sometimes also called a social network; for example, as shown in Fig. 11.1. Note that as such arcs indicate that interaction takes place, and interaction in principle means that persons affect each other, from a dynamical perspective they can also be considered relations that represent mutual causal effects on certain mental states of persons.

The area of network models to describe social interactions has already a longer tradition (see also Chap. 1, Sect. 1.4). It has gradually developed stronger and at the same time a similar development concerning the concept ‘network’ took place in other disciplines as well, such as Biology, Neuroscience, Mathematics, Physics, Economics, Informatics, Artificial Intelligence, and Web Science; see, for example Boccaletti et al. (2006), Valente (2010), Giles (2012). A unified perspective on networks as a generic modeling concept was developed which forms the basis of the Network-Oriented Modeling perspective used here.

To characterize different types of network models, over the years some criteria have been identified:

- The length of the *shortest path* between two given nodes in the network
- The *number of connections of a given node* (called the *degree*), and how that varies over the network
- The occurrence of *clusters* of tightly connected groups of nodes (*sub-communities*)
- The variation of *strengths of the connections* over the network

In this chapter dynamics in relation to social interactions will be addressed. Two main types of dynamics are distinguished: dynamics based on a given structure of interactions (modeled by a non-adaptive network), and dynamics of the social interactions (modeled by an adaptive network). In the former case the network model stays the same, but *states* (nodes) in the network may change their *activation level* over time. In the latter case the *connections* within the network model may change, for example, connections can be added or removed, or their *weights* may increase or decrease.

In this chapter it will be discussed how adaptive network models can be made for social interaction processes. Persons will simply be modeled as nodes in a network indicating for each person just one state concerning, for example, an emotion, a belief, an opinion, or a behaviour. For a network modeling social interaction processes again conceptual representations of two types are possible, as shown in Chap. 2: a graphical representation or a matrix representation, and numerical representations that can be used to perform simulation experiments (see Chap. 2).

In this chapter, first an overview of different types of network models is presented: small world networks and random networks (Sect. 11.2), scale-free networks (Sect. 11.3), and weighted networks with variation in connection strengths (Sect. 11.4). After this in Sect. 11.5 a brief overview is given of the different types of dynamics for networks modeling social interaction processes. In Sect. 11.6 the case of dynamics based on social contagion is discussed, modeled by a non-adaptive temporal-causal network. After that in a number of sections different types of dynamics of a social interaction structure are addressed, modeled by adaptive temporal-causal networks: based on the ‘birds of a feather flock together’ (homophily) principle (Sect. 11.7), based on the ‘more becomes more’ principle (Sect. 11.8), and based on the ‘interaction connects’ principle (Sect. 11.9). Finally Sect. 11.10 is a discussion.

11.2 Small World Networks and Random Networks

The first characteristic of a network considered concerns paths in the network. Examples of the paths in a network are the following. In the World Wide Web consider a Web page is connected to another one when a link to this other Webpage occurs. Then a path from one Webpage to another one is based on a number of mouse clicks bringing you there. The shortest path from one Webpage to another one is the minimal number of mouse clicks needed to get from the one to the other. For another example, consider the connections between persons for which are considered connected when they know each other: they are acquaintances or more (e.g., at least once they have spoken with each other). Suppose a k -th order contact is a contact with shortest path of length k . You can wonder, for example, which order of contact you have with Barack Obama. In general it may be expected that the shortest paths between two arbitrary nodes in a larger network may still be very long, unless the nodes in the network have many connections. For example, it may be expected that the number of steps needed to connect you to an arbitrary person anywhere in the world can be very large, as every person knows only a limited

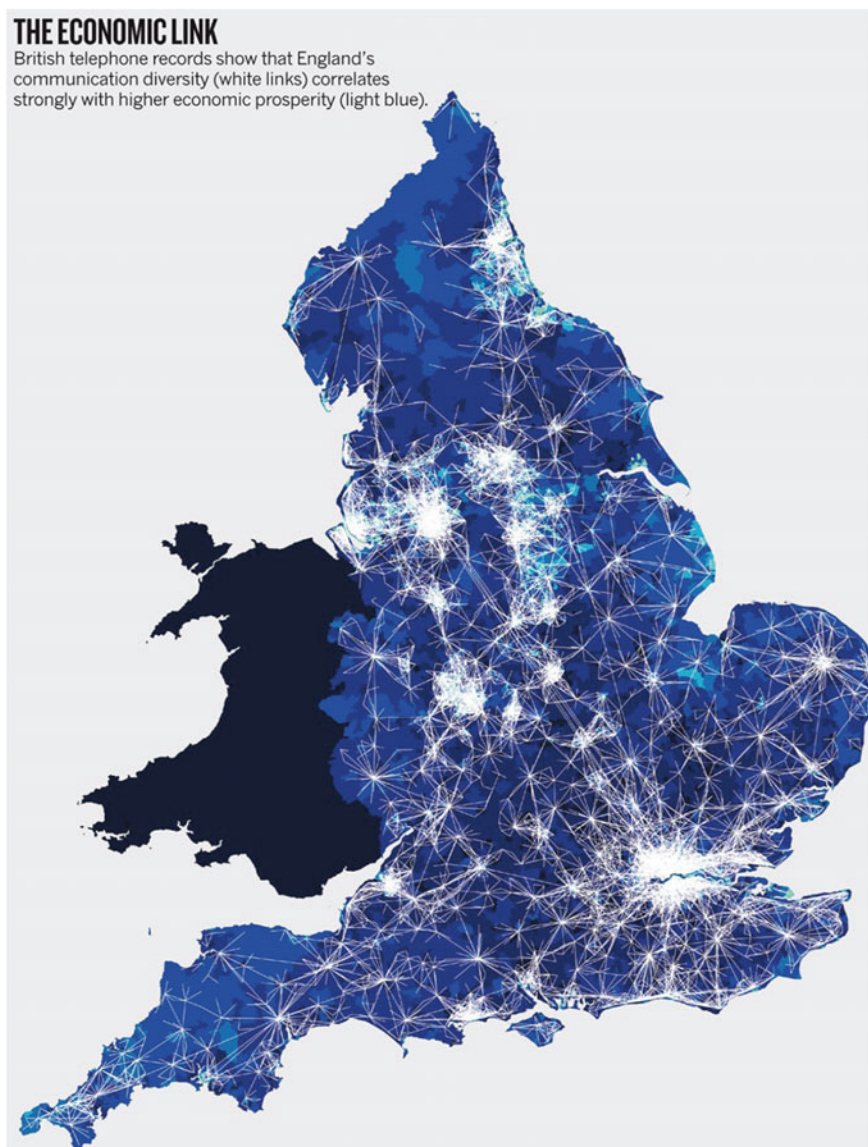


Fig. 11.1 Example social network based on telephone calls, adopted from Giles (2012)

number of other persons. However, in the real world shortest paths seem much shorter than one would expect. The notion ‘small world’ is sometimes used in relation to the experience when you find out that a person you meet for the first time knows somebody you know as well; this would make $k = 2$. Usually this is felt as a surprise and not seldom expressed by something like ‘What a small world!’.

11.2.1 *Small World Networks*

The term *small world network* has been adopted to indicate a network in which

- the average number of (direct) connections per node is low
- but still the shortest paths are short

Of course, when practically all nodes in a network are directly connected, and therefore the number of connections per node is high, then automatically the shortest paths will be very short, and even 1 for a fully connected network (everybody knows everybody). It has been found that many networks in the real world that have been analyzed fall in the class of small world networks. This does not only concern networks describing social interactions, but also, for example, in Biology networks describing chains of metabolic reactions, in Neuroscience networks describing processes in the brain, and in Web Science networks describing the World Wide Web. For the human society the conjecture has even been put forward that any two persons on earth are in a k -th order contact with k at most 6.

11.2.2 *Random Networks*

Random networks satisfy the criterion that there is a fixed number p in the interval $[0, 1]$ such that for every pair of nodes they are connected with probability p . Figure 11.2 shows examples of random networks for $p = 0.1$, 0.25 , and 0.5 , respectively.

In the case of a random network with n nodes the number of connections per node, also called *degree*, has a Poisson distribution with top around $p(n - 1)$ and going to zero on both sides: around 1 and around n . More precisely, suppose the fraction of nodes with degree k is denoted as $P(k)$. The graph of $P(k)$ as a function of k showing the (frequency or probability) distribution of the degrees over a network has its top at $p(n - 1)$. For example, for a random network of $n = 20$ nodes with $p = 0.5$, this top is at $p(n - 1) = 9.5$. Note that for $p = 1$ (fully connected network) there are no nodes with degree lower than $n - 1$: in that case all nodes have the maximal degree $n - 1$.

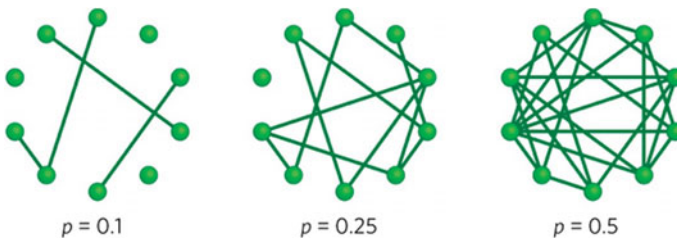


Fig. 11.2 Examples of random networks for different link probabilities; adopted from Perseguers et al. (2010)

11.3 Distribution of Node Degrees and Scale-Free Networks

Networks describing processes in the real world usually have different shapes of distributions of degrees, with more nodes in the lower degrees and less in the higher degrees, as is the case for the small world networks. So, the notion of a random network does not describe at all any arbitrary network describing processes that occur in the real world. Instead, the study of networks describing real world processes has led to the notion of *scale-free networks*.

11.3.1 Scale-Free Networks

A *scale-free network* is defined as a network that satisfies a distribution of node degrees according to a *power law*: the number or fraction of nodes with degree k , is approximated by a function

$$f(k) = c k^{-\gamma}$$

for some exponent γ which typically (but not always) may be between 2 and 4, and c a constant. Three patterns for such a distribution with different exponents are shown in Fig. 11.3. This clearly is totally different from the distribution for a random network.

The name ‘scale-free’ was derived from a property of the function f representing the power law. This function f satisfies the property that for any α a β exists such that for all x it holds: $f(\alpha x) = \beta f(x)$. Therefore scaling up the degree k by any factor α still provides the same shape of distribution. The conjecture is that many networks

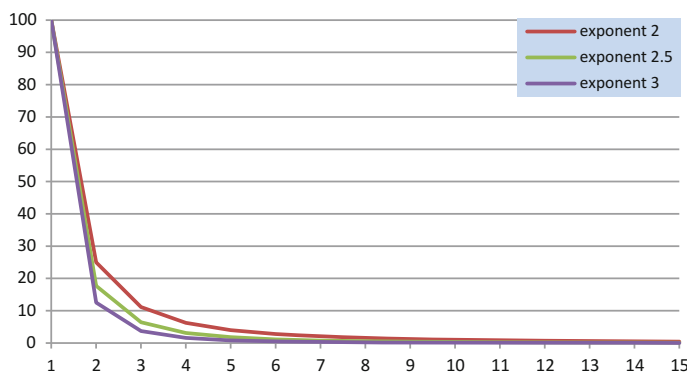


Fig. 11.3 Power law distribution of node degrees for a scale-free network for $c = 100$ and exponent $\gamma = 2, 2.5$ and 3 , respectively

describing processes in the real world are scale-free, for example, networks based on World Wide Web links, networks based on chemical reaction chains in biological networks, and networks based on social interaction.

In Barabasi and Albert (1999, p. 510) the results shown in Table 11.1 are put forward. For a large network based on WWW pages it turned out that it satisfies a power law with γ about 2.1. Also a network based on actors with connections when they were playing at least once in the same movies had a power law structure, with γ about 2.3. Similarly it was found that the network based on citations between scientific papers satisfies a power law with γ about 3 and an electrical power grid (in the western US) with γ about 4. Such empirical results have led to the hypothesis that networks in practice usually are scale-free with a distribution of degrees according to a power law with γ between 2 and 4. Further empirical evidence for this hypothesis is still being acquired.

11.3.2 Identifying a Power Law

As a simple example, suppose the network shown in Fig. 11.4 is given. The question may be posed whether this network is scale-free, and if so, which exponent γ applies.

By counting, or by the matrix representation, the degree distribution can be found as shown in the middle column of Table 11.2.

This distribution has been approximated by a power law distribution function $f(k) = c k^{-\gamma}$ by choosing $\gamma = 2$ and $c = 18$; see the column at the right hand side in Table 11.2. In Fig. 11.5 for both (empirical and power law) distributions the graph

Table 11.1 Some empirical results showing for which exponent γ the power law holds for networks in practice

Network	WWW	Actors	Citations	Power grid
γ	2.1	2.3	3	4

Fig. 11.4 Small example of a scale-free network (adopted from https://en.wikipedia.org/wiki/Scale-free_network)

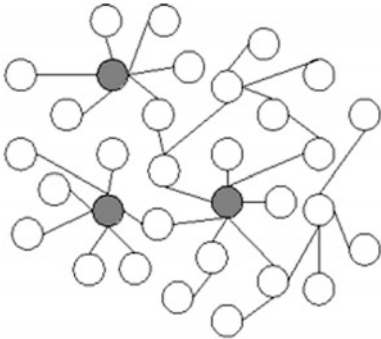
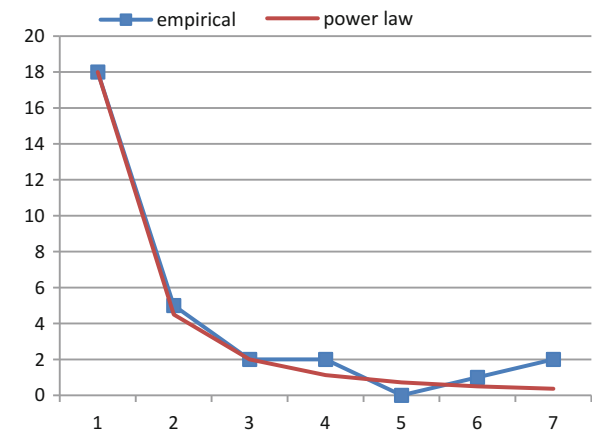


Table 11.2 Empirical data on degrees for the network in Fig. 11.4 compared to a power function with $\gamma = 2$ and $c = 18$

Degree	Empirical	Power law
1	18	18.00
2	5	4.50
3	2	2.00
4	2	1.13
5	0	0.72
6	1	0.50
7	2	0.37

Fig. 11.5 Graph for empirical degree data and power law for the network in Fig. 11.10



is shown. This shows that the network depicted in Fig. 11.4 is scale-free with γ around 2.

A useful approach to estimate which exponent γ fits well to a set of empirical data is by applying the natural logarithm **log** to the power law, thus obtaining:

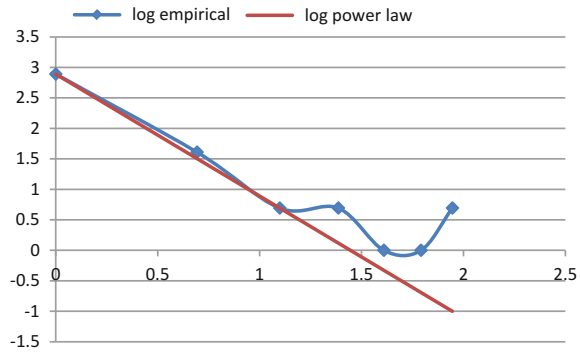
$$\log f(k) = \log c - \gamma \log k$$

This is a linear relation between **log** $f(k)$ and **log** k , which will show as a straight line in a graph. When this logarithm is also applied to the empirical data, the data shown Table 11.3 are obtained by applying the logarithm to the cells of Table 11.2.

Table 11.3 Logarithm of empirical data on degrees for the network in Fig. 11.10 compared with a logarithm of a power function (with $\gamma = 2$ and $c = 18$)

Log degree	Log empirical	Log power law
0	2.89037176	2.890372
0.693147	1.60943791	1.088189
1.098612	0.69314718	0.03398
1.386294	0.69314718	-0.71399
1.609438	$-\infty$	-1.29417
1.791759	0	-1.7682
1.94591	0.69314718	-2.16899

Fig. 11.6 Double logarithmic graph for empirical degree data and power law for the network in Fig. 11.4



This can be depicted in a (double logarithmic) graph as shown in Fig. 11.6. Note that the value $-\infty$ resulting from applying **log** to 0 is depicted here as 0.

In this form for any given set of empirical data the parameters c and γ can be estimated by minimizing the sum of the squares of the distances, as is done by a linear least square approximation method; see also Clauset et al. (2009).

11.3.3 Clusters and Bridges

In scale-free networks, often a form of clustering structure can be found; see, for example, Figs. 11.1 and 11.4. Sometimes also clusters themselves again have their own clustering structure, which gives the network a repetitive structure. The clusters are identified as subgroups of nodes that have a relatively high number of mutual connections but practically no connection with nodes outside the cluster. If there are no such connections between a cluster and other clusters at all, the whole cluster would be isolated. However, usually there are at least some connections from one cluster to another; they are called *bridge connections* or simply *bridges*. A special case is when all nodes within a cluster are mutually connected, in which case the cluster sometimes is called a *clique*.

11.4 Weak Ties, Strong Ties and Weighted Connections

Above it was assumed that nodes in a network are connected or not connected. However, it has since long been recognized that connections can have different gradual strengths and types. Some can be very weak, like being acquaintances, others can be very strong like between best friends or in a partner or family relation. Weak connections are often called *weak ties*, and strong connections *strong ties*; e.g., see Fig. 11.7. In general, a *weighted network* is a network in which for each pair of nodes a *connection* or *tie weight* (or strength) is defined as a number in the

interval $[0, 1]$. The weak ties are those with lower weight and strong ties those with higher weight.

A claim made by Rapoport (1957) is the *weak tie hypothesis*: if A has strong ties to both B and C, then there is a high probability that B and C at least have a weak tie. As an example, your two best friends probably at least know each other, or, as another example, your parents and partner probably know each other (e.g., Fig. 11.8).

The idea is (Granovetter 1973, 1983) that strong ties usually occur within the clusters and lead to weak ties making the cluster more tightly connected. Moreover, the bridges between clusters are usually formed by weak ties, and are essential for contagion from one cluster to another; this principle is sometimes indicated by *the power of weak ties*.

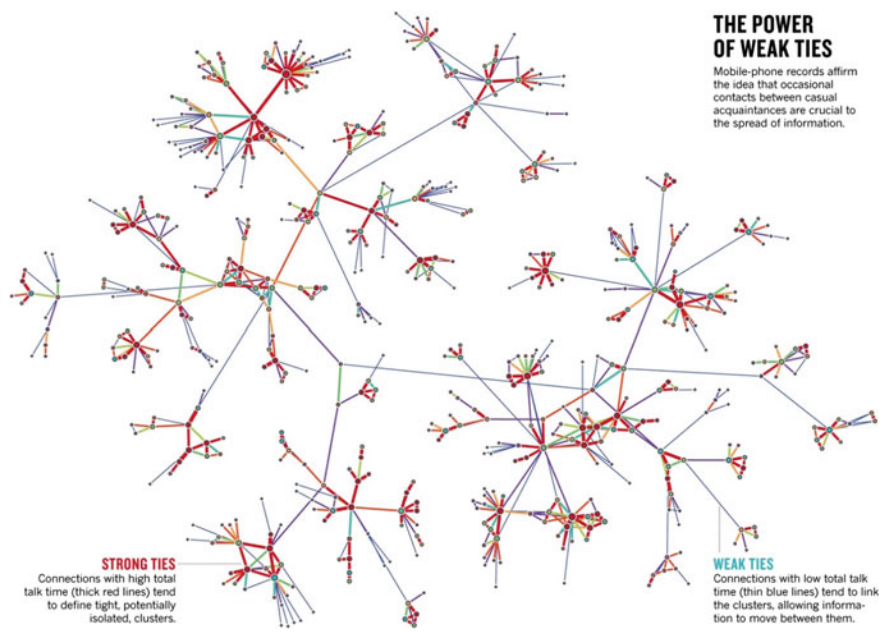
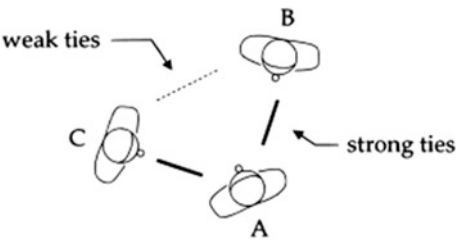


Fig. 11.7 Weak and strong ties based on telephone calls; adopted from Giles (2012)

Fig. 11.8 Weak tie hypothesis (adopted from http://en.wikipedia.org/wiki/Interpersonal_ties)



By Krackhardt (1992) the *strong tie hypothesis* was formulated. This hypothesis claims that sometimes for more fundamental types of contagion weak ties are not sufficient. Especially for contagion that comes with uncertainty a connection is needed that makes a person feel secure, which usually is a characteristic of a strong tie but not of a weak tie. Such more fundamental types of contagion would not make it over bridges when they are only weak ties.

For weighted networks the matrix representation contains nonnegative real numbers as (weight) values, usually taken from the interval $[0, 1]$. An example of this is shown in Table 11.4 (the 0's have been left out), which is a matrix representation of the network depicted in graphical representation in Fig. 11.9.

Note that the connections in this example network are not bidirectional.

Based on these weights as an approximation the influence of a given node on the network can be estimated, and also conversely, the influence of the network on this node. For example, node E has 7 outgoing connections, so this node has probably much influence on the network. In the other direction, for example, node K has 5 incoming connections; this is an indication that this node receives much direct influence from the network. These exerted and received influences can be estimated numerically in a more precise manner by taking the sums of the weights of the outgoing and the incoming connections, respectively. In Table 11.5 these sums have been added to the matrix from Table 11.4. In Table 11.5 the rightmost column depicts for each node the overall influence exerted in the network directly. It can be seen that node E indeed has one of the highest numbers in this column, but node A and J still have higher numbers, so they have more direct influence. Similarly, in Table 11.5 in the bottom row the numbers of overall received direct influence are depicted. It can be seen here that node K has a high number, but in this row node E has the highest number, so that node will be influenced most in the network.

The question can be posed whether this approximation obtained by adding up the connection weights is adequate. It may be the case that a node has much influence on a number of connected nodes, but these nodes in turn have not much influence. Such a situation would limit the value of the calculations made in Table 11.5. In fact, the approach shown in Table 11.5 can be considered a first-order approximation of the influence from and the influence on a node in a network. Also higher order approximations are possible, for example, taking into

Fig. 11.9 Example network represented by the matrix in Table 11.4

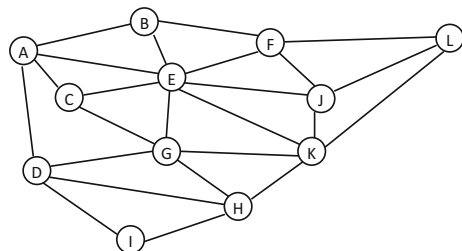


Table 11.4 Matrix representation of a weighted network

From	To											
	A	B	C	D	E	F	G	H	I	J	K	L
A		0.6	0.8	1	0.5							
B	0.7				0.4	0.8						
C	0.5				1		0.9					
D	0.6						0.4	0.8	0.7			
E	0.3	0.5	0.4			0.3	0.1			0.7	0.3	
F		0.6			0.3					0.7		0.2
G			1	0.2	0.3			0.1			0.2	
H				0.5			0.4		0.8		0.5	
I				0.9				1				
J					0.4	0.7					0.9	0.7
K					0.3		0.2	0.3		0.1		0.2
L						0.9				0.8	0.9	

Table 11.5 Overall exerted and received direct influence in a network

From	To												
	A	B	C	D	E	F	G	H	I	J	K	L	Directly exerted influence
A		0.6	0.8	1	0.5								2.9
B	0.7				0.4	0.8							1.9
C	0.5				1		0.9						2.4
D	0.6						0.4	0.8	0.7				2.5
E	0.3	0.5	0.4			0.3	0.1			0.7	0.3		2.6
F		0.6			0.3					0.7		0.2	1.8
G			1	0.2	0.3			0.1			0.2		1.8
H				0.5			0.4		0.8		0.5		2.2
I				0.9				1					1.9
J					0.4	0.7					0.9	0.7	2.7
K					0.3		0.2	0.3		0.1		0.2	1.1
L						0.9				0.8	0.9		2.6
Directly received influence	2.1	1.7	2.2	2.6	3.2	2.7	2	2.2	1.5	2.3	2.8	1.1	

account the influence of the influenced nodes. For example, for node E a second order approximation can be calculated as follows. Take the nodes influenced by node E and multiply for each X of them the weight of the connection from E to node X by the first-order influence number of node X, and the add all these products (Table 11.6):

Table 11.6 Estimating second-order influence

	A	B	C	D	E	F	G	H	I	J	K	L	2nd order influence of E
E's outgoing weights	0.3	0.5	0.4			0.3	0.1			0.7	0.3		5.72
1st order influences of nodes connected from E	2.9	1.9	2.4			1.8	1.8			2.7	1.1		
Products of the above two numbers	0.87	0.95	0.96			0.54	0.18			1.89	0.33		

This results in a number of 5.72 for node E’s second-order influence. Similarly, third or higher order influences can be calculated; all these numbers will provide better approximations. It can be useful to compare such estimations with outcomes of simulation experiments with (social) contagion in such a network.

11.5 Different Types of Dynamics in Networks Based on Social Interaction

In this section a number of types of dynamics in relation to networks based on social interaction are discussed. Given any network, two main types of dynamics can be addressed: dynamics *within* the given network, and dynamics *of* the network. In the former case the network stays the same, but *states* (nodes) in the network may change their *activation level* over time. In the latter case the network *connections* change, for example their *weights* may increase or decrease, which makes it an adaptive network. Both types of dynamics often also occur together.

The first type of dynamics is usually indicated by words such as *diffusion*, *spread* or *contagion*. The node states that are spread can be of different types varying from communicating by body language, oral verbal communication, and written verbal communication to communication by transferring chemicals (e.g., Groot et al. 2012). For example:

- *Information, beliefs, opinions or valuations*
Examples:
 - Did you know that A is dating B?
 - There is a new restaurant nearby
 - Do you also like this movie?
 - This phone is not good

- *Emotions*

Examples:

- Showing a happy face, talking with a sad voice,
- Saying or writing that you feel bad
- Spreading your sweat and pheromones (e.g., for fear or disgust, or for sexual arousal)

- *Diseases*

Examples:

- Contagion of your flu to somebody else, or even to your whole network
- Tuberculosis epidemic

- *Behaviour, habits or lifestyle*

Examples:

- Contagion of sport activities among friends
- Smoking behaviour
- Use of alcohol or drugs

The second type of dynamics concerns changes in the connectivity structure of the network: it becomes an *adaptive network*. Also such changes can be of different types, such as:

- *Link addition*

Examples:

- Meeting somebody,
- Becoming friends with somebody
- Citing a paper

- *Link removal*

Examples:

- Unfriend somebody
- Breaking up a relationship

- *Strengthening a link*

Examples:

- After more and more activities together a weak tie becomes a strong tie
- Discovering that you have much in common with somebody

- *Weakening a link*

Examples:

- Weakening of a friendship due to moving to an area far away
- Neglecting the contact with somebody during a very busy period

A number of general principles are used in networks in order to conduct simulation experiments and mathematical analysis. Examples of such principles are (e.g., Newman 2003):

- *Random link addition*

When it is assumed that persons meet each other by chance, this can be modeled by link addition in a network between randomly chosen pairs of nodes.

- *Triadic closure*

Examples of this principle are *transitive addition* (if A is connected to B and B to C, then A will get a connection with C: ‘your friends are my friends’), and the *weak tie hypothesis* discussed in Sect. 11.4.

For these adaptive networks three different examples will be addressed: based on the homophily principle, the more becomes more principle, and the interaction connects principle describing the effect of the frequency and intensity of actual interaction on connections.

- *Homophily*

This principle indicates that the more similar the states of two connected nodes are, the stronger their connection will become: ‘birds of a feather flock together’; see, for example, Byrne (1986), McPherson et al. (2001), Mislove et al. (2010). When also the states are assumed dynamic, this principle can be combined with contagion of states into a circular causal relation:

$$\text{state} \rightleftharpoons \text{link}$$

When in a practical situation a network is found in which similar states and strong connections occur together, due to such a circular causal relation it is difficult to tell which of the two principles (or both?) was causing this situation; see also, for example Aral et al. (2009), Shalizi and Thomas (2011), Steglich et al. (2010), Mundt et al. (2012).

- *More becomes more*

This principle expresses that nodes that already have more connections get more extra connections than nodes with less connections (the rich become more rich and the poor remain poor). Analyses have been made showing that applying this principle usually leads to scale-free networks; e.g., Price (1976), Barabási and Albert (1999), Newman (2003).

- *Interaction connects*

The third approach to adaptive networks addresses how the weights of connections are affected by the actual interaction taking place. This is based on the assumption that the more frequent and the more intense the interaction between two persons, the stronger they will become related; see, e.g., Hove and Risen (2009), Pearce et al. (2015).

In Sect. 11.6 a basic temporal-causal network model is discussed for social contagion within a given, non-adaptive network. Sections 11.7–11.9 address different types of adaptive networks, based on different principles. In Sect. 11.7 an adaptive temporal-causal network model based on the *homophily principle* is discussed. According to this principle connections strengthen more when the persons are more similar in their state (the more you are alike, the more you like each other). In Sect. 11.8 an adaptive network based on the *more becomes more principle* is 17 discussed. This principle assumes that persons that already have more and stronger connections also attract more and stronger (new) connections, more than persons with less and weaker connections (more popular becomes still more popular). Finally, in Sect. 11.9 it is discussed how dynamics of network connections can be modeled based on information about actual interaction between persons, both in the sense of frequency and of intensity: *interaction connects principle*.

11.6 Social Contagion

In a social context often persons affect each other in a number of ways. These mutual influences can concern emotions, but also opinions and beliefs, and different types of behaviour. Such processes are usually called *social contagion*, and specific types of it such as social contagion of emotions are called emotion contagion. From a behavioural perspective social contagion processes can be modeled as shown by a conceptual representation in Fig. 11.10 (in this case for two persons only). Here the (executed) behaviour *b* (for example, showing a face expression with a specific emotion) of one person is sensed by another person and vice versa. Within a given person this behaviour can be affected by some stimulus *s*, but also by sensing the same behaviour shown by another person.

Here the weights of the connections between the persons are assumed constant, thus defining the given network. However, the network itself can be adaptive in the

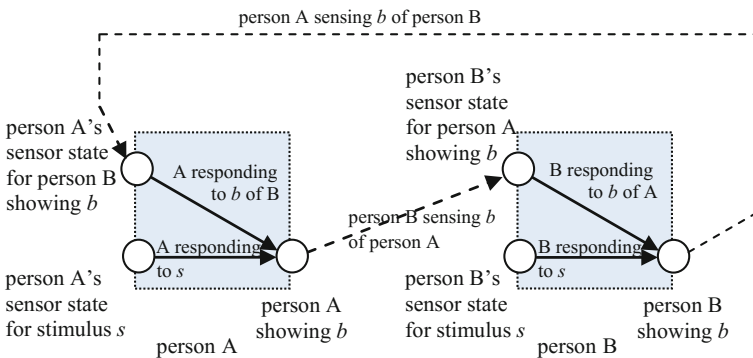


Fig. 11.10 Graphical conceptual representation of a temporal-causal network model for a behavioural view on social contagion

sense that the connections between the persons are dynamic too. In Sects. 11.7–11.9 this is discussed in some more detail.

The graphical conceptual representation for this temporal-causal network model for this behavioural view on social contagion can be transformed into a numerical representation as shown in Box 11.1. Here execution states for b of a person Y are named by $es_{Y,b}$, and sensor states of a person Y for sensing the execution state $es_{X,b}$ of person X are denoted by $ss_{Y,es_{X,b}}$. The connection weight $\omega_{Y,sensing_b_of_X}$ represents the aspects of the interaction and relation between person X and Y that determine how well b of X is observed by Y . For example, when b is a face expression of X and this is not observed at all by Y , because they have contact with a very low frequency and often only by text messages, this weight value is (close to) 0. However, when the connection is through social media, often certain expression means can be used to reflect one's face expression (emojicons); this enables the connection weight being higher.

The connection weight $\omega_{Y,responding_to_b_of_X}$ represents how responsive Y is for b of X , after Y has sensed b of X . Persons and connections may show strong differences in this. Note that the connections $\omega_{Y,responding_to_b_of_X}$ and $\omega_{Y,responding_to_s}$ determine some balance between the effect of the stimulus s in the response in comparison to the effect of the observed behaviour from others. Persons Y with high $\omega_{Y,responding_to_b_of_X}$ compared to $\omega_{Y,responding_to_s}$ will easily adapt to other persons, whereas persons Y with high $\omega_{Y,responding_to_s}$ compared to $\omega_{Y,responding_to_b_of_X}$ will be more difficult to affect by others.

Be aware that in this model the overall contagion effect of the execution state $es_{Y,b}$ of person Y on the execution state $es_{X,b}$ of person X is achieved in two steps: first sensing with weight $\omega_{Y,sensing_b_of_X}$ and then responding with weight $\omega_{Y,responding_to_b_of_X}$. Together these weights determine how strong this contagion effect is.

For persons X and Y with $X = A$ and $Y = B$, or $X = B$ and $Y = A$:

LP1 Person Y sensing b of person X

$$\begin{aligned} dss_{Y,es_{X,b}}/dt &= \eta_{ss_{Y,es_{X,b}}} [c_{ss_{Y,es_{X,b}}} (\omega_{Y,sensing_b_of_X} es_{X,b}) - ss_{Y,es_{X,b}}] \\ ss_{Y,es_{X,b}}(t + \Delta t) &= ss_{Y,es_{X,b}}(t) + \eta_{ss_{Y,es_{X,b}}} [c_{ss_{Y,es_{X,b}}} (\omega_{Y,sensing_b_of_X} es_{X,b}(t)) - ss_{Y,es_{X,b}}(t)] \Delta t \end{aligned}$$

LP2 Person Y responding to stimulus s and to b of person X

$$\begin{aligned} des_{Y,b}/dt &= \eta_{es_{Y,b}} [c_{es_{Y,b}} (\omega_{Y,responding_to_s} ss_{Y,s}, \omega_{Y,responding_to_b_of_X} ss_{Y,es_{X,b}}) - es_{Y,b}] \\ es_{Y,b}(t + \Delta t) &= ps_{Y,b}(t) + \eta_{ps_{Y,b}} [c_{ps_{Y,b}} (\omega_{Y,responding_to_s} ss_{Y,s}(t), \\ &\quad \omega_{Y,responding_to_b_of_X} ss_{Y,es_{X,b}}(t)) - es_{Y,b}(t)] \Delta t \end{aligned}$$

The symbols are explained as follows:

s	external stimulus
b	behavioural response, for example, an emotion
$es_{Y,b}$	execution state: person Y shows response b
$ss_{Y,Z}$	sensor state: person Y senses Z with $Z = s$ or $Z = es_{X,b}$
η_S	update speed parameter for state S
$c_S(..)$	combination function for state S

Box 11.1 Numerical representation of a temporal-causal network model for a behavioural view on social contagion between two persons

Note that the above model covers only the interaction between two persons. In a network model for social interaction, the effect of multiple persons on a given person has to be aggregated, so this is still to be addressed. Moreover, in the abstract behavioural perspective shown above, each person is characterised by three states: one behaviour execution state and two sensor states. In the area of networks models for social interaction and social media, often a simplified and still more abstract view is used, in which only one state per person is modeled, which can be considered as the behaviour shown to others. This simplification makes it easier to handle large numbers of persons in such a network, up to hundreds, or thousands, or even millions.

As an illustration, a more abstract temporal-causal network model is described with contagion effects of multiple persons on each person and based on each person Y 's state $es_{Y,b}$. First a conceptual model is shown in Fig. 11.11. Note that more connections may occur than the ones drawn. For example, for every connection from an person X to an person Y there may also exist a connection back from Y to X .

This conceptual representation can be transformed into a numerical representation of the model as follows; for all Y in the network the differential and difference equation

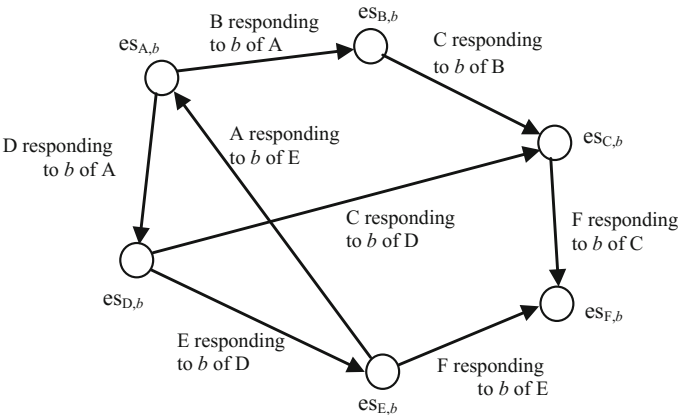


Fig. 11.11 Conceptual representation for simple temporal-causal network model for social contagion

are shown in Box 11.2. Here for each connection from an person X to a person Y there is a direct connection from the execution state $es_{X,b}$ of person X to the execution state $es_{Y,b}$ of person Y , with weight indicated by $\omega_{Y_responding_to_X,b}$. Note that this is different from the model shown in Box 11.1 above, where there is a two step process, first from execution state of person X to sensor state of person Y , and then from sensor state of person Y to execution state of person Y . In that model $\omega_{Y_responding_to_b_of_X}$ is indicating only the weight of the internal connection from sensor state to execution state of person Y , whereas $\omega_{Y_sensing_b_of_X}$ indicates the weight of the external connection from execution state of person X to sensor state of person Y . So, the connection weight parameter $\omega_{Y_responding_to_X,b}$ in the model in Box 11.2 relates to the two parameters $\omega_{Y_sensing_b_of_X}$ and $\omega_{Y_responding_to_b_of_X}$

in the model in Box 11.1. In a simple case with two persons and identity combination functions it holds that the former weight can be set as the product of the latter two weights:

$$\omega_{Y_responding_to_X,b} = \omega_{Y_sensing_b_of_X} * \omega_{Y_responding_to_b_of_X}$$

For all persons X_i and Y in $\{A, \dots, F\}$ with X_i the persons connected toward Y :

LP1 Person Y responding by b

$$\begin{aligned} des_{Y,b}/dt &= \eta_{es_{Y,b}} [c_{es_{Y,b}} (\omega_{Y_responding_to_X_1,b} es_{X_1,b}, \dots, \\ &\quad \omega_{Y_responding_to_X_n,b} es_{X_n,b}) - es_{Y,b}] \\ es_{Y,b}(t + \Delta t) &= es_{Y,b}(t) + \eta_{es_{Y,b}} [c_{es_{Y,b}} (\omega_{Y_responding_to_X_1,b} es_{X_1,b}(t), \dots, \\ &\quad \omega_{Y_responding_to_X_n,b} es_{X_n,b}(t)) - es_{Y,b}(t)] \Delta t \end{aligned}$$

The symbols are explained as follows:

b	behavioural response, for example, an expressed emotion
$es_{Y,b}$	execution state: person Y shows response b
$\eta_{es_{Y,b}}$	update speed parameter for state $es_{Y,b}$
$c_{es_{Y,b}}(\cdot)$	combination function for state $es_{Y,b}$
$\omega_{Y_responding_to_X,b}$	weight of connection from $es_{X,b}$ to $es_{Y,b}$

Box 11.2 Numerical representation of a simple temporal-causal network model for a behavioural view on social contagion of multiple persons

An often used combination function for aggregation of contagion effects of multiple persons is the *scaled sum function*:

$$c_{es_{Y,b}}(V_1, \dots, V_n) = \mathbf{ssum}_{\omega_{Y,b}}(V_1, \dots, V_n) = (V_1 + \dots + V_n) / \omega_{Y,b}$$

where the scaling factor $\omega_{Y,b}$ is the sum of the weights of connections from all others X_i for $i = 1, \dots, n$ to Y :

$$\omega_{Y,b} = \omega_{Y_responding_to_X_i,b} + \dots + \omega_{Y_responding_to_X_n,b}$$

This combination makes that the aggregated impact from other persons is a weighted average of the individual impacts with weights $\omega_{Y_responding_to_X_i,b}/\omega_{Y,b}$ proportional to the connection weights $\omega_{Y_responding_to_X_i,b}$ and with sum of them 1:

$$\omega_{Y_responding_to_X_1,b}/\omega_{Y,b} + \dots + \omega_{Y_responding_to_X_n,b}/\omega_{Y,b} = 1$$

and

$$\begin{aligned} c_{es_{Y,b}}(\omega_{Y_responding_to_X_1,b} es_{X_1,b}(t), \dots, \omega_{Y_responding_to_X_n,b} es_{X_n,b}(t)) \\ &= \text{ssum}_{\omega_{Y,b}}(\omega_{Y_responding_to_X_1,b} es_{X_1,b}(t), \dots, \omega_{Y_responding_to_X_n,b} es_{X_n,b}(t)) \\ &= (\omega_{Y_responding_to_X_1,b} es_{X_1,b}(t) + \dots + \omega_{Y_responding_to_X_n,b} es_{X_n,b}(t))/\omega_{Y,b} \\ &= (\omega_{Y_responding_to_X_1,b}/\omega_{Y,b}) es_{X_1,b}(t) + \dots + (\omega_{Y_responding_to_X_n,b}/\omega_{Y,b}) es_{X_n,b}(t) \end{aligned}$$

The interpretation of this combination function is that a person Y adapts the level of his or her state $es_{Y,b}$ to what is observed by Y as the average level of the state $es_{X,b}$ in the network; this can result in adapting to the majority in the network. However, if the connections only relate to a small and relatively isolated part of the network, the average over this part can deviate from the average of the network as a whole.

For modeling dynamics within networks models for social interaction the simplest idea is that one state X for each person is considered, and denoted by X_A for person A . This state X can be any type of state, either internal or externally observable, for example, an internal state of feeling an emotion, or an expressed emotion state, or an (internal) intention, or an action performed, or a belief or opinion. The connection weights are denoted as $\omega_{A,B}$ for the connection from person A to person B . The model from Box 11.2 reformulated in this more abstract form is shown in Box 11.3. For some further variants of temporal-causal network models for social contagion, see Bosse et al. (2015).

LP1 Social contagion of state X in a social network

$$dX_B/dt = \eta_B[c_B(\omega_{A_1,B}X_{A_1}, \dots, \omega_{A_k,B}X_{A_k}) - X_B]$$

$$X_B(t + \Delta t) = X_B(t) + \eta_B[c_B(\omega_{A_1,B}X_{A_1}(t), \dots, \omega_{A_k,B}X_{A_k}(t)) - X_B(t)]\Delta t$$

The symbols are explained as follows:

X_B	person B 's state X
η_B	update speed parameter for state X_B
$c_B(..)$	combination function for state X_B
$\omega_{A,B}$	weight of connection from A to B

Box 11.3 Numerical representation of a model for a simple network view on social contagion

11.7 Adaptive Network Dynamics and the Homophily Principle

In a simple model the network characteristics are specified by the connection strengths $\omega_{A,B}$. When a simple model for an adaptive network is to be designed, a first question to answer is how these connection strengths are changing, and, in particular, the other states affecting them have to be identified. The simple model with adaptive network connections discussed here is obtained as a refinement of the simple temporal-causal network model discussed in Sect. 11.6, Box 11.3. A first choice made is to assume that the connection strengths $\omega_{A,B}$ are affected by the activation levels of the connected states of A and B . Such a dependency is depicted in Fig. 11.12. Note that by adding these effects on the connection strengths cyclic relationships occur; for example:

$$X_B \text{ affects } \omega_{A,B} \quad \text{and in turn} \quad \omega_{A,B} \text{ affects } X_B$$

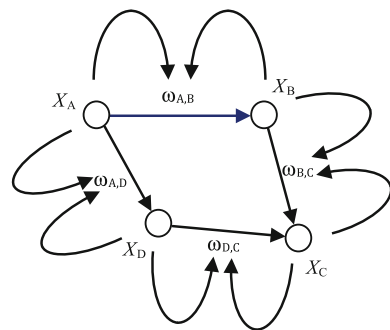
Also note that at a conceptual graphical level there is similarity with modeling Hebbian learning for connections between states within a person (see Chap. 2, Sect. 2.10, Fig. 2.24). However, the two types of models differ, due to different types of combination functions used to describe the impacts of the different states on the connections.

A next step is to determine how exactly the connection strengths are affected by the activation levels. This is needed to obtain a dynamic equation for $\omega_{A,B}$. For the current model the dynamic $\omega_{A,B}$ are assumed to change over time based on a principle similar to the one from Parunak et al. (2011): the closer the activation levels of the states of the interacting persons, the stronger the mutual connections between the persons will become, and the higher the difference between the activation levels, the weaker they will become. In other words:

- activation levels close to each other imply a strong upward change in $\omega_{A,B}$
- activation levels far apart imply a downward change of $\omega_{A,B}$

Sometimes this is called a *homophily principle*: the more you are alike, the more you like (each other); for example, see Byrne (1986), McPherson et al. (2001),

Fig. 11.12 Conceptual representation of an adaptive temporal-causal network model for the homophily principle



Mislove et al. (2010). Another way of formulating this principle is: birds of a feather flock together.

As an example of this principle in practical use, if you wonder whether there is a chance to become connected with somebody, you might consider whether you often like and agree on the same things. Some may even consider to fake such a similarity (but of course this is a doubtful strategy as it may seriously harm how your unique identity is perceived). It can often be observed that persons that have close relationships or friendships are alike in some respects. For example, they go to the same clubs, take the same drinks, have the same opinions, vote for the same or similar parties. Such observations might be considered support for the homophily principle: in the past they were attracted to each other due to being alike. However, also a different explanation is possible: they were together and due to that they affected each other's states by social contagion, and therefore they became alike. So, the cyclic relation between X_B and $\omega_{A,B}$ mentioned above leads to two possible causal explanations of a state of being alike and being connected:

$$\begin{array}{ll} \text{being connected} & \rightarrow \text{being alike} \\ \text{being alike} & \rightarrow \text{being connected} \end{array}$$

Such circular causal relations make it difficult to determine what came first. It may be a state just emerging from a cyclic process without a single cause. For more discussion on this issue, for example, see Aral et al. (2009), Shalizi and Thomas (2011), Steglich et al. (2010), Mundt et al. (2012).

The homophily principle may be formalised numerically using the following general format and a combination function $c_{A,B}(V_1, V_2, W)$ that still has to be determined:

$$\begin{aligned} \omega_{A,B}(t + \Delta t) &= \omega_{A,B}(t) + \eta_{A,B}[c_{A,B}(X_A(t), X_B(t), \omega_{A,B}(t)) - \omega_{A,B}(t)]\Delta t \\ \mathbf{d}\omega_{A,B}/\mathbf{d}t &= \eta_{A,B}[c_{A,B}(X_A, X_B, \omega_{A,B}) - \omega_{A,B}] \end{aligned}$$

Here it is assumed that the values of $\omega_{A,B}$ stay within the interval $[0, 1]$ and in particular the conditions

$$c_{A,B}(V_1, V_2, 0) \geq 0 \quad \text{and} \quad c_{A,B}(V_1, V_2, 1) \leq 1$$

are fulfilled. The combination function $c_{A,B}(\dots)$ is assumed to depend on the one hand on W and on the other hand on the difference $|V_1 - V_2|$ (which always is between 0 and 1) in such a way that lower values of $|V_1 - V_2|$ relate to higher values of $c_{A,B}(V_1, V_2, W)$, and higher values of $|V_1 - V_2|$ relate to lower values of $c_{A,B}(V_1, V_2, W)$; in particular:

$$\begin{aligned} |V_1 - V_2| \text{ high} &\Rightarrow \omega_{A,B} \text{ decreasing} \Rightarrow \mathbf{d}\omega_{A,B}(t)/\mathbf{d}t \leq 0 \Rightarrow c_{A,B}(V_1, V_2, W) \leq W \\ |V_1 - V_2| \text{ low} &\Rightarrow \omega_{A,B} \text{ increasing} \Rightarrow \mathbf{d}\omega_{A,B}(t)/\mathbf{d}t \geq 0 \Rightarrow c_{A,B}(V_1, V_2, W) \geq W \end{aligned}$$

In particular this holds for the extreme values 0 and 1 of $|V_1 - V_2|$:

$$\begin{aligned} |V_1 - V_2| = 1 &\Rightarrow c_{A,B}(V_1, V_2, W) \leq W \\ |V_1 - V_2| = 0 &\Rightarrow c_{A,B}(V_1, V_2, W) \geq W \end{aligned}$$

Moreover, to model the homophily principle discussed above, a reasonable assumption is that there is a monotonically decreasing relation:

the higher $|V_1 - V_2|$, the lower $c_{A,B}(V_1, V_2, W)$

Furthermore, it can be assumed that it only depends on this difference $|V_1 - V_2|$ and not on the values of V_1 and V_2 themselves. Then as a simplification the combination function $c_{A,B}(\dots)$ can be expressed as a function $h_{A,B}(D, W)$ of $D = |V_1 - V_2|$ and W :

$$c_{A,B}(V_1, V_2, W) = h_{A,B}(D, W)$$

As discussed above, the function $h_{A,B}$ is assumed to be monotonically decreasing in D :

$$\begin{aligned} D_1 \leq D_2 &\Rightarrow h_{A,B}(D_1, W) \geq h_{A,B}(D_2, W) \\ D \text{ high} &\Rightarrow h_{A,B}(D, W) \leq W \\ D \text{ low} &\Rightarrow h_{A,B}(D, W) \geq W \end{aligned}$$

and in particular for the extreme values 0 and 1 of D it holds $h_{A,B}(1, W) \leq W$ and $h_{A,B}(0, W) \geq W$. Moreover,

$$h_{A,B}(D, 0) \geq 0 \quad \text{and} \quad h_{A,B}(D, 1) \leq 1$$

Somewhere between low values of $D = |V_1 - V_2|$ (with $h_{A,B}(D, W) \geq W$) and high values of D (with $h_{A,B}(D, W) \leq W$) there can be assumed a value for D for which $h_{A,B}(D, W) = W$. This is called the *threshold value*, indicated by $\tau_{A,B}$; so

$$\begin{aligned} h_{A,B}(D, W) &\geq W \quad \text{when } D \leq \tau_{A,B} \\ h_{A,B}(D, W) &= W \quad \text{when } D = \tau_{A,B} \\ h_{A,B}(D, W) &\leq W \quad \text{when } D \geq \tau_{A,B} \end{aligned}$$

So, for this threshold value $\tau_{A,B}$ it holds:

- an upward change of connection weight $\omega_{A,B}$ occurs when $|V_1 - V_2| < \tau_{A,B}$
- no change of connection weight $\omega_{A,B}$ occurs when $|V_1 - V_2| = \tau_{A,B}$
- a downward change of connection weight $\omega_{A,B}$ occurs when $|V_1 - V_2| > \tau_{A,B}$

A very simple example of a continuous function $h_{A,B}(D, W)$ satisfying the above conditions for a given value of W is obtained when the threshold value $\tau_{A,B}$ is assumed to be fixed; then a simple decreasing linear function in D through the point

with coordinates $(\tau_{A,B}, W)$ can be used (i.e., through the point with $D = \tau_{A,B}$ and $h_{A,B}(D, W) = W$):

$$h_{A,B}(D, W) = W + \alpha(\tau_{A,B} - D)$$

for some α , that still can be chosen. This example may have the disadvantage that when $W = 0$ it may be negative (when $D > \tau_{A,B}$) or when $W = 1$ it may be higher than 1 (when $D < \tau_{A,B}$): then it does not fulfil the conditions

$$h_{A,B}(D, 0) \geq 0 \quad \text{and} \quad h_{A,B}(D, 1) \leq 1$$

and this may make the weight value $\omega_{A,B}$ go outside the interval $[0, 1]$. This can be remedied by choosing α as a function $\alpha(W)$ of W which can suppress the term $\tau_{A,B} - D$ when W comes closer to 0 or 1:

$$h_{A,B}(D, W) = W + \alpha(W)(\tau_{A,B} - D)$$

When this function $\alpha(W)$ is assumed to be always ≥ 0 and close to 0 when W is close to 0 or 1, then this can keep the value of $\omega_{A,B}$ within the interval $[0, 1]$. This can be satisfied by the function

$$\alpha(W) = W(1 - W)$$

which is 0 for $W = 0$ and for $W = 1$, and positive between these values with a maximum 0.25 for $W = 0.5$. This makes that ω is changing slowly in the neighbourhood of 0 or 1, thus achieving that ω does not cross these boundaries. Then the following example function fulfilling the above conditions is obtained:

$$h_{A,B}(D, W) = W + W(1 - W)(\tau_{A,B} - D)$$

As a variant of this, the following function can be obtained as a quadratic function of D :

$$h_{A,B}(D, W) = W + W(1 - W)(\tau_{A,B}^2 - D^2)$$

Yet another variant can be defined using a logistic function:

$$h_{A,B}(D, W) = W + W(1 - W)(0.5 - 1/(1 + e^{-\sigma_{A,B}(D - \tau_{A,B})}))$$

In Fig. 11.13 these linear, quadratic and logistic functions are depicted, with D on the horizontal axis and $h_{A,B}(D, W)$ on the vertical axis for $\tau_{A,B} = 0.2$ and three fixed values for W (0.1, 0.5, 0.9); the depicted logistic variant has steepness $\sigma_{A,B} = 6$.

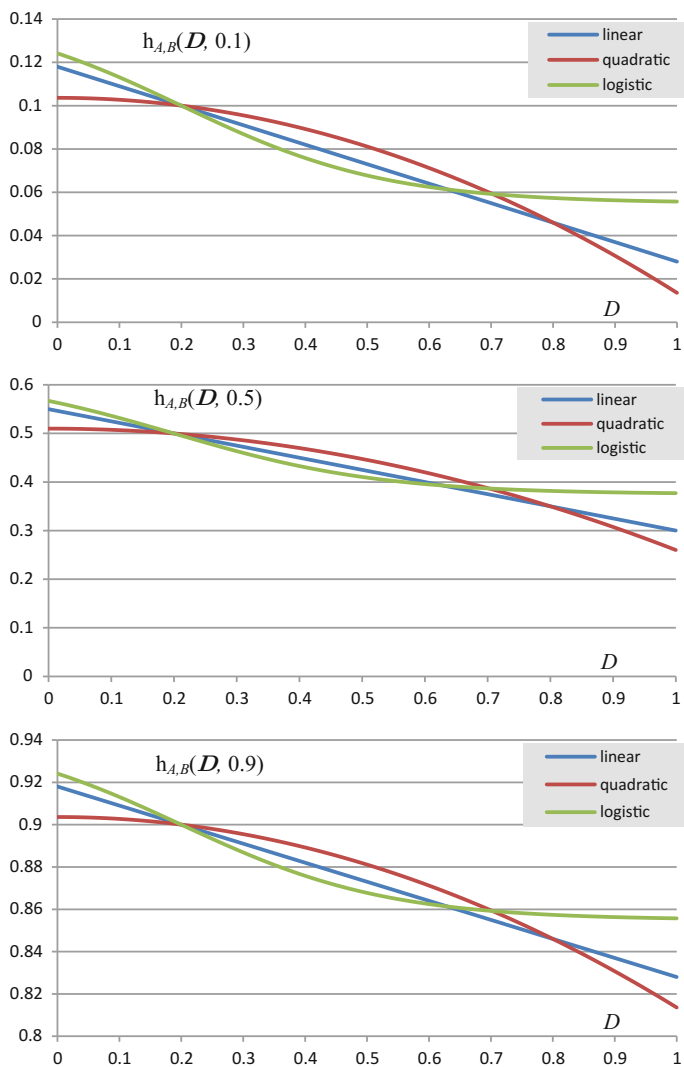


Fig. 11.13 Linear, quadratic and logistic form of the function $h_{A,B}(D, W)$ for the change of $\omega_{A,B}$ for values 0.1, 0.5 and 0.9 for W , respectively, and $\tau_{A,B} = 0.2$; steepness is 6

For the combination function $c_{A,B}(V_1, V_2, W)$ the above choices for $h_{A,B}(D, W)$ translate into:

$$c_{A,B}(V_1, V_2, W) = W + W(1 - W) (\tau_{A,B} - |V_1 - V_2|)$$

$$c_{A,B}(V_1, V_2, W) = W + W(1 - W) (\tau_{A,B}^2 - (V_1 - V_2)^2)$$

$$c_{A,B}(V_1, V_2, W) = W + W(1 - W) (0.5 - 1/(1 + e^{-\sigma_{A,B}(|V_1 - V_2| - \tau_{A,B})}))$$

Using these functions, the dynamic relations for $\omega_{A,B}$ are, respectively, as shown in Box 11.4.

For all persons A and B in the social network:

$$\begin{aligned} d\omega_{A,B}/dt &= \eta_{A,B} \omega_{A,B} (1 - \omega_{A,B}) (\tau_{A,B} - |X_A - X_B|) \\ \omega_{A,B}(t + \Delta t) &= \omega_{A,B}(t) \\ &\quad + \eta_{A,B} \omega_{A,B}(t) (1 - \omega_{A,B}(t)) (\tau_{A,B} - |X_A(t) - X_B(t)|) \Delta t \\ d\omega_{A,B}/dt &= \eta_{A,B} \omega_{A,B} (1 - \omega_{A,B}) (\tau_{A,B}^2 - (X_A - X_B)^2) \\ \omega_{A,B}(t + \Delta t) &= \omega_{A,B}(t) \\ &\quad + \eta_{A,B} \omega_{A,B}(t) (1 - \omega_{A,B}(t)) (\tau_{A,B}^2 - (X_A(t) - X_B(t))^2) \Delta t \\ d\omega_{A,B}/dt &= \eta_{A,B} \omega_{A,B} (1 - \omega_{A,B}) (0.5 - 1/(1 + e^{-\sigma_{A,B}(|X_A - X_B| - \tau_{A,B})})) \\ \omega_{A,B}(t + \Delta t) &= \omega_{A,B}(t) \\ &\quad + \eta_{A,B} \omega_{A,B}(t) (1 - \omega_{A,B}(t)) (0.5 - 1/(1 + e^{-\sigma_{A,B}(|X_A(t) - X_B(t)| - \tau_{A,B})})) \Delta t \end{aligned}$$

The symbols are explained as follows:

X_A, X_B	state of person A resp. B
$\omega_{A,B}$	weight of the connection from person A to person B
$\eta_{A,B}$	update speed parameter for the connection from person A to person B
$\tau_{A,B}$	threshold for connection adaptation

Box 11.4 Numerical representation of an adaptive temporal-causal network model for social dynamics based on the homophily principle: linear, quadratic and logistic variant

In simulations it turns out that in principle all $\omega_{A,B}$ converge to either 0 or 1, and the levels of the states may converge in different emerging clusters of persons, where the connections between persons within one cluster converge to 1 and the connections between persons in different clusters converge to 0; for example, see Fig. 11.14

Three similar but slightly more complex variants can be made, of which the quadratic variant is described in Sharpanskykh and Treur (2014):

$$\begin{aligned} h_{A,B}(D, W) &= W + \text{Pos}(\eta_{A,B}(\tau_{A,B} - D)) (1 - W) - \text{Pos}(-\eta_{A,B}(\tau_{A,B} - D)) W \\ h_{A,B}(D, W) &= W + \text{Pos}(\eta_{A,B}(\tau_{A,B}^2 - D^2)) (1 - W) - \text{Pos}(-\eta_{A,B}(\tau_{A,B}^2 - D^2)) W \\ h_{A,B}(D, W) &= W + \text{Pos}(\eta_{A,B}(0.5 - 1/(1 + e^{-\sigma_{A,B}(D - \tau_{A,B})}))) (1 - W) \\ &\quad - \text{Pos}(-\eta_{A,B}(0.5 - 1/(1 + e^{-\sigma_{A,B}(D - \tau_{A,B})}))) W \end{aligned}$$

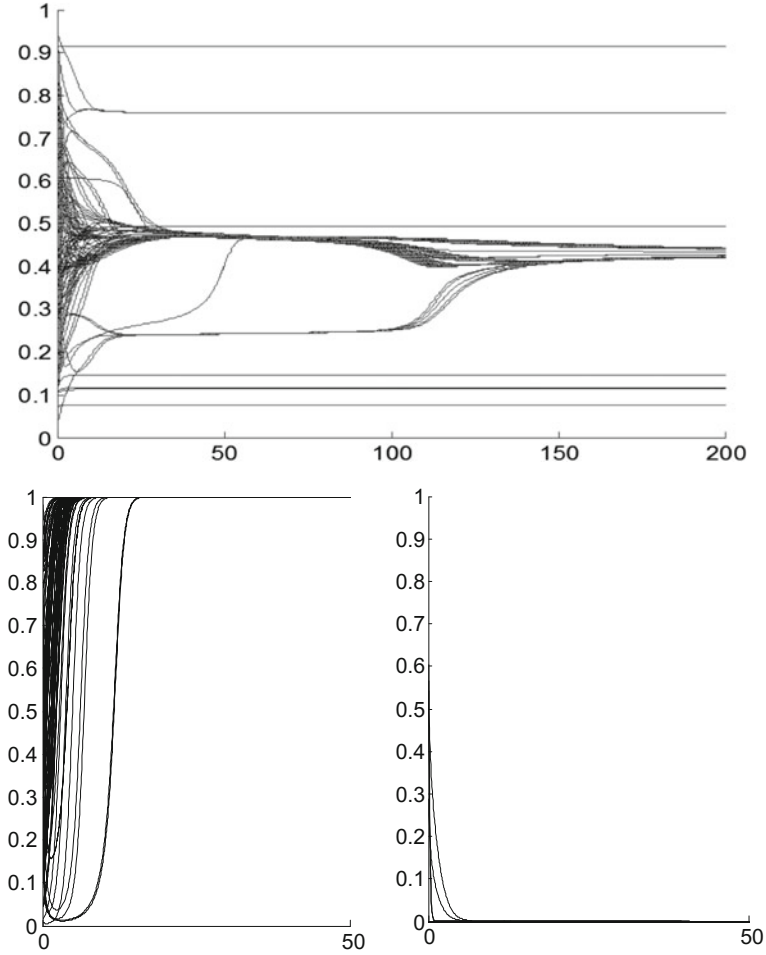


Fig. 11.14 Emergence of different clusters in an adaptive temporal-causal network model based on the homophily principle (model depicted in Box 11.4) (*upper* state levels, *lower left* connections within clusters; *lower right* connections between nodes in different clusters)

Here $\text{Pos}(x) = (|x| + x)/2$, which returns x when x is positive and 0 when x is negative. These models make that the approaching of the boundaries 0 and 1 of the interval $[0, 1]$ of ω is slow, thus making ω not crossing these boundaries, but ω departing from the neighbourhood of these boundaries is not slow. In Sharpanskykh and Treur (2014) example simulations can be found using the second, quadratic model.

11.8 Adaptive Networks and the More Becomes More Principle

Another type of model for a dynamic connection from a person B to A takes into account to which extent other persons C connect to person A . The idea behind this is that somebody who is very popular seems worth connecting to. Sometimes this is called the ‘more becomes more’ principle. For example, if B is followed by many others C on Twitter, then B seems to be interesting to follow for A as well. As the connections of others to B may change over time, this will imply that also A will have a dynamic connection to B , and in turn this connection will affect the connection of others to B over time as well. This can be modeled taking into account the weights $\omega_{C_i,B}$ for $i = 1, \dots, k$ of all connections from others C_i to B as shown in Box 11.5 and Fig. 11.15. Here $c_{A,B}(\dots)$ is a combination function for the values $\omega_{C_1,B}, \dots, \omega_{C_k,B}$, for example, a logistic sum function, or a scaled sum function with scale factor the number of other persons. Note that the latter combination function only takes into account the average strengths of the connections, not the total number of them.

LP1 Network connections affected by other connections

$$d\omega_{A,B}/dt = \eta_{A,B}[c_{A,B}(\omega_{C_1,B}, \dots, \omega_{C_k,B}) - \omega_{A,B}]\Delta t$$

$$\omega_{A,B}(t + \Delta t) = \omega_{A,B}(t) + \eta_{A,B}[c_{A,B}(\omega_{C_1,B}(t), \dots, \omega_{C_k,B}(t)) - \omega_{A,B}(t)]\Delta t$$

The symbols are explained as follows:

$\omega_{A,B}$ weight of connection from A to B

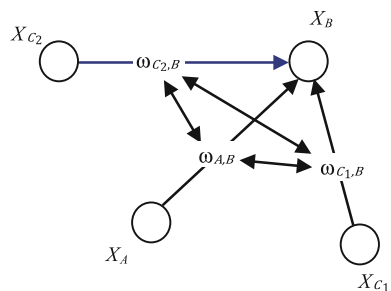
$\eta_{A,B}$ update speed parameter for $\omega_{A,B}$

$c_{A,B}(\dots)$ combination function for $\omega_{A,B}$

Box 11.5 Numerical representation of an adaptive temporal-causal network model for dynamics of connections based on the more becomes more principle

Note that a network modeling the initiation of connections is not automatically a network indicating social contagion; this will depend on the application considered. For example, a network modeling a connection from A to B when A is following

Fig. 11.15 Conceptual representation of an adaptive temporal-causal network model for the more becomes more principle

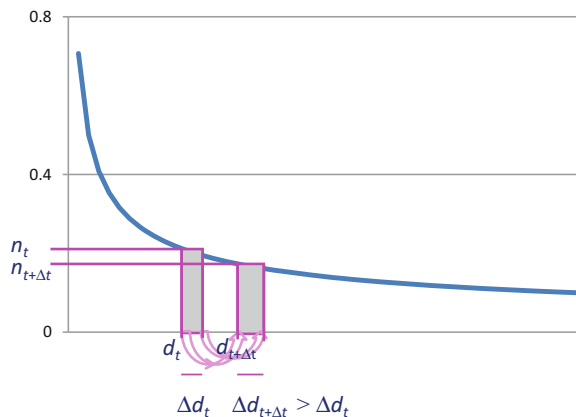


B on Twitter will not play a role in social contagion from A to B . For social contagion the opposite network plays a role where a connection from A to B occurs when A is followed by B , which is not initiated by A but by B : on Twitter and most other social media you cannot appoint your own followers. As another example, when A often contacts B for advice, and this advice is often taken over by A , then the initiation connection is from A to B but the contagion connection is from B to A . In other cases it may be different. For example, if A wants to announce an event or new product, he or she can choose an occasion where many others will see the message, for example, posting it on a suitable forum; in such a case both the initiation and the social contagion are directed from A to the others.

The ‘more becomes more’ principle has also been used to provide an explanation for the empirical evidence that most real-world networks are scale-free. The idea is that the typical distribution of degrees according to a power law emerges from an evolving network when it is assumed that the network dynamics is based on a ‘more become more’ principle; see, for example, Barabási and Albert (1999), Krapivsky et al. (2000), Krapivsky and Redner (2001, 2003). An indication of the type of argument followed is illustrated in Fig. 11.16. Here the distribution of nodes (vertical axis) over degrees (horizontal axis) is depicted; this distribution is assumed stable over time. Some time point t is considered and the focus is at the nodes with some degree d_t at t (see at the horizontal axis). There is a (relative) number or density n_t of them (vertical axis). Moreover, the nodes with degree between d_t and a bit higher $d_t + \Delta d_t$ are considered, an interval of length Δd_t at the horizontal axis. The (relative) number of nodes with degree within this interval is represented in Fig. 11.16 by the area of the (left) rectangle above that interval. This area is approximated by $n_t \Delta d_t$.

Now consider a time step from t to $t + \Delta t$. Due to growth of the number of connections, the nodes with degree d_t at time t will have a higher degree $d_{t+\Delta t}$ at $t + \Delta t$, and the nodes with degree $d_t + \Delta d_t$ at time t will have a higher degree $d_{t+\Delta t} + \Delta d_{t+\Delta t}$ at $t + \Delta t$. Due to the ‘more becomes more’ principle, from $d_t < d_t + \Delta d_t$ it follows that from t to $t + \Delta t$ the nodes with degree $d_t + \Delta d_t$ at time t will

Fig. 11.16 Emerging scale-free network from an adaptive network



get more new connections than the nodes with degree d_t at time t . Therefore the increase in degree of these nodes with degree $d_t + \Delta d_t$ at time t will be higher:

$$\Delta d_{t+\Delta t} > \Delta d_t$$

The numbers of nodes previously represented at t by the left rectangle are represented at $t + \Delta t$ by the right rectangle. Moreover, because they describe the same nodes, the areas indicated as shaded are the same:

$$n_t \Delta d_t = n_{t+\Delta t} \Delta d_{t+\Delta t}$$

Given this equality, from $\Delta d_{t+\Delta t} > \Delta d_t$ (more becomes more principle) it follows that $n_{t+\Delta t} < n_t$. Therefore the distribution is monotonically decreasing. By a more complex argument it has been derived that based on some more precise assumptions on the formalisation of the more becomes more principle, a distribution is obtained according to a power law; for example, see Barabási and Albert (1999), Krapivsky et al. (2000), Krapivsky and Redner (2001, 2003) for more details.

11.9 Adaptive Networks and Actual Interaction Over Time

In the next variant of a model with adaptive network characteristics, a different assumption is taken as a point of departure for how the connection weights $\omega_{A,B}$ are changing. In contrast to the previous models in which the change of connection weights was assumed dependent on the activation levels of the connected states or of the weights of other connections, here they are assumed to be affected by the actual interaction taking place. The idea behind this is the principle that the more interaction (or joint activities) you actually have with somebody, the stronger you will become related; e.g., Hove and Risen (2009), Pearce et al. (2015). An example of this *interaction connects principle* in practical use is: if you want to become connected with somebody, consider to follow the strategy to often undertake activities with this person. The dependencies according to this general principle are depicted in Fig. 11.9. Here the actual amount of interaction from person X to person Y concerning state S is indicated by ai_{XY} .

Studies discuss how connection ('tie') strength in network models for both online and offline social interaction relates to (i) interaction frequency, (ii) emotional intensity of content, and (iii) emotional support and closeness (Gilbert and Karahalios 2009; Granovetter 1983; Marsden and Campbell 1990). The number of questions asked in a message also relates to connection strength. For example, in Morris et al. (2010) it was found that many participants' questions in online social interaction were answered by friends they rated as close and that closeness of a friendship was a motivator to answer questions. At least part of this literature uses this relation between connection strength and actual interaction to formulate an

operational, measurable definition of connection strength. In that case it concerns a definitional relationship between connection strength and interaction, in contrast to, for example, literature such as Hove and Risen (2009), Pearce et al. (2015) in which a causal relationship is assumed. In the model described below some relationship between connection strength and interaction is assumed, but it is not necessarily assumed to be definitional. Given the network shown in Fig. 11.17, a numerical representation for the model of the dynamics of the connection strength is modeled as shown in Box 11.4.

LP1 Network connections affected by actual interaction

$$d\omega_{A,B}/dt = \eta_{\omega_{A,B}} [c_{\omega_{A,B}}(ai_{A,B}) - \omega_{A,B}] \Delta t$$

$$\omega_{A,B}(t + \Delta t) = \omega_{A,B}(t) + \eta_{\omega_{A,B}} [c_{\omega_{A,B}}(ai_{A,B}(t)) - \omega_{A,B}(t)] \Delta t$$

The symbols are explained as follows:

A, B	any two persons
$\omega_{A,B}$	weight of connection from A to B
$ai_{A,B}$	strength of actual interaction from A to B
$\eta_{\omega_{A,B}}$	update speed parameter for $\omega_{A,B}$
$c_{\omega_{A,B}}(\dots)$	combination function for $\omega_{A,B}$

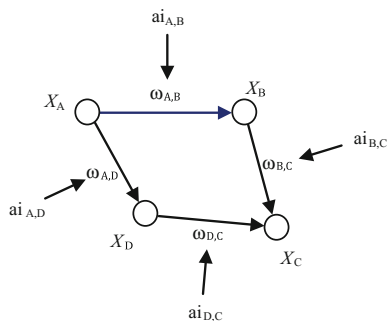
Box 11.6 Numerical representation of an adaptive temporal-causal network model for dynamics of a social interaction affected by actual interaction

Here $c_{\omega_{A,B}}(\dots)$ is some combination function which, for example, may be the identity function or multiplication by a fixed number or a logistic function. For a similar model, see van Breda et al. (2012).

Using quantity and quality of actual interaction

In the model described in Fig. 11.17, it is assumed that the actual amounts of interaction can be obtained in an independent manner, for example, by monitoring the different occurrences of interaction in a social medium such as Twitter. Such monitoring can follow a differentiated approach by taking into account two different (both measurable) aspects of interaction: quantity of the interaction (measurable as

Fig. 11.17 Conceptual representation of an adaptive temporal-causal network model with dynamic connection strengths based on actual interaction



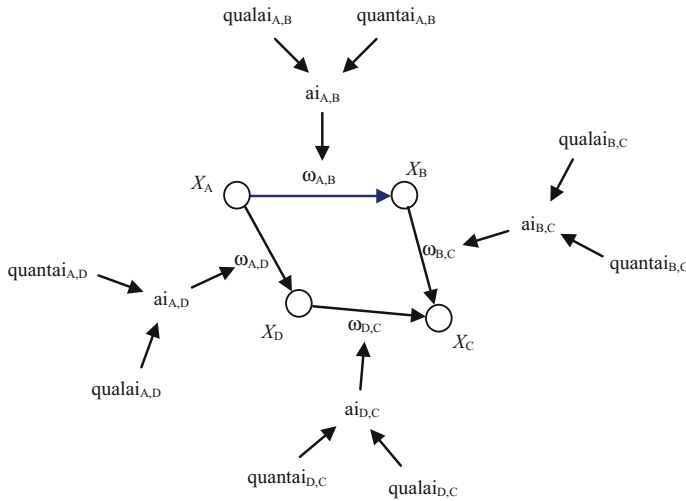


Fig. 11.18 Conceptual representation of a model of dynamic connection weights based on quality and quantity of actual interaction

frequency and duration), but also the quality or intensity of the interaction, which, for example, may relate to the (measurable) level of emotion and number of questions in the interaction. A graphical conceptual representation for such a model is shown in Fig. 11.18.

The next step is to determine a numerical representation of this model for the dynamics of the connection strengths. For the amount of interaction $ai_{X,Y}$ the product of the quantity and quality measures is taken:

$$ai_{X,Y} = quantai_{X,Y} * qualai_{X,Y}$$

Moreover, for the quantity a frequency measure is defined as a function g of the frequency (number of messages per time unit, which is a day) with values in the interval $[0, 1]$:

$$quantai_{A,B} = g(\text{frequency})$$

for example, with 4 considered as maximal frequency:

$$g(V) = \begin{cases} V/4 & \text{when } V \leq 4 \\ 1 & \text{when } V > 4 \end{cases}$$

Quality is defined as a weighted average of the emotion levels expressed in the messages and a value in $[0, 1]$ for the average number of questions in the messages during one time unit:

$$\text{qualai}_{X,Y} = w_1 * \text{average emotion level} + w_2 * \text{average question level}$$

So the connection impact becomes

$$\text{ai}_{X,Y} = g(\text{frequency}) (w_1 * \text{average emotion level} + w_2 * \text{average question level})$$

Using mutual effects between connection strength and actual interaction

Another variant of the model described above can be made if it is taken into account that as an opposite influence the weight of a connection also affects the quality and quantity of the interaction. This is shown in Fig. 11.19 as an extension of what is depicted in Fig. 11.18. Such a model expresses that as soon as a connection of a certain strength exists, this is confirming and maintaining itself as due to the existence of this connection interaction takes place through this connection. This can be modeled in a detailed manner by putting

$$d\text{quantai}_{A,B}/dt = \eta_{qt_{A,B}} [c_{qt_{A,B}}(\omega_{A,B}) - \text{quantai}_{A,B}]$$

$$d\text{qualai}_{A,B}/dt = \eta_{ql_{A,B}} [c_{ql_{A,B}}(\omega_{A,B}) - \text{qualai}_{A,B}] \Delta t$$

$$\text{quantai}_{A,B}(t + \Delta t) = \text{quantai}_{A,B}(t) + \eta_{qt_{A,B}} [c_{qt_{A,B}}(\omega_{A,B}(t)) - \text{quantai}_{A,B}(t)] \Delta t$$

$$\text{qualai}_{A,B}(t + \Delta t) = \text{qualai}_{A,B}(t) + \eta_{ql_{A,B}} [c_{ql_{A,B}}(\omega_{A,B}(t)) - \text{qualai}_{A,B}(t)] \Delta t$$

Here $c_{qt_{A,B}}(\dots)$ is some combination function, for example, multiplication by some factor, or a logistic function.

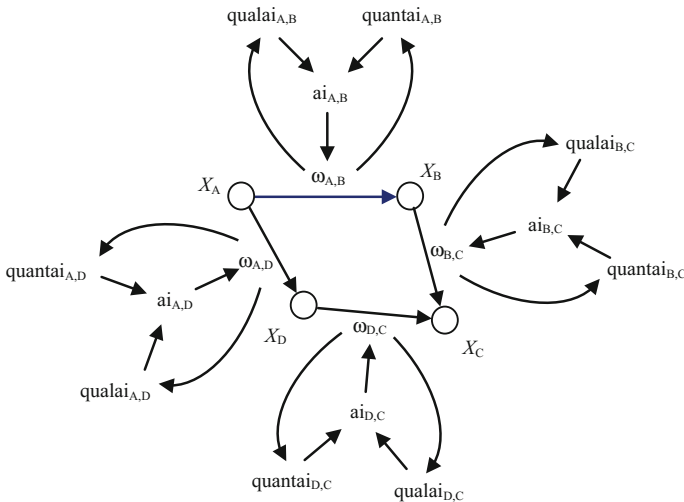


Fig. 11.19 Mutually affecting dynamic connection strengths and actual interaction

11.10 Discussion

In this chapter, an overview of different types of networks describing social interaction was presented, varying from small world networks, random networks, scale-free networks to weighted networks in which connections have different weights. Moreover, different types of dynamics for network models for social interaction have been addressed in some detail, covering both dynamics within a given network based on social contagion and dynamics of an adaptive network, for which the network structure changes. For the latter type of (adaptive) dynamics three different principles were addressed: the homophily principle, the more becomes more principle, and the interaction connects principle. The model for the homophily principle is based on Sharpanskykh and Treur (2014); the model for the interaction connects principle was inspired by van Breda et al. (2012). It has been shown how the Network-Oriented Modeling approach based on temporal-causal network models as described in Chap. 2 applies well to the domain of social interaction.

As an extension of the work described in this chapter, an integrative adaptive temporal-causal network model has been designed by integrating models for social contagion and both the adaptive network models for the homophily principle and the more becomes more principle as discussed above; see Blankendaal et al. (2016). This integrative model has been validated for an empirical data set involving development of relations between teenagers.

An important role of social interaction is that they form a basis for diffusion or contagion processes for various matters, for example, diseases, information, innovations, opinions, emotions, behaviours, lifestyles. Monitoring and analysing the dynamics of given diffusion or contagion processes is one thing. Many dynamic models have been put forward to support such analysis, varying from agent-based to population-based models. However, having such analysis means available, they can be used for *prediction* as well, and also for *what-if simulation*: predicting what will happen if some action is undertaken. More generally, methods can be developed to determine what types of *network intervention* actions can be undertaken under which conditions in order to achieve some specific goal. Examples of such goals are:

- avoiding that an epidemic will develop
- achieving that many persons will know about a new product that you bring out
- achieving that more people will adopt a healthy lifestyle
- achieving that elderly persons are not in a situation in which they have almost no social contacts

As an example, persons with large numbers of connections may be identified (for example, with thousands or even millions of followers on Twitter), as they may play a crucial role to reach many nodes. As another example, in a strongly clustered kind of network mutual influence or contagion can take place according to a kind of repeated sequence of waterfalls, where at each step some time is passing to get a

cluster affected, after which a next cluster is affected. In such a process the bridge connections play an important role, so network interventions may focus on them. In Valente (2010, 2012) more can be found on this area of network interventions.

References

- S. Aral, L. Muchnik, A. Sundararajan, Distinguishing influence based contagion from homophily driven diffusion in dynamic networks. *Proc. Natl. Acad. Sci. (USA)* **106**(2), 1544–1549 (2009)
- A.L. Barabási, R. Albert, Emergence of scaling in random networks. *Science* **286**, 509–512 (1999)
- R. Blankendaal, S. Parinussa, J. Treur, A temporal-causal modelling approach to integrated contagion and network change in social networks, in *Proceedings of the 22nd European Conference on Artificial Intelligence, ECAI'16* (IOS Press, 2016)
- S. Boccaletti, V. Latorab, Y. Morenod, M. Chavez, D.-U. Hwanga, Complex networks: structure and dynamics. *Phys. Rep.* **424**(2006), 175–308 (2006)
- T. Bosse, R. Duell, Z.A. Memon, J. Treur, C.N. van der Wal, Agent-based modelling of emotion contagion in groups. *Cogn. Comput. J.* **7**(2015), 111–136 (2015)
- D. Byrne, The attraction hypothesis: do similar attitudes affect anything? *J. Personal. Soc. Psychol.* **51**(6), 1167–1170 (1986)
- A. Clauset, C.R. Shalizi, M.E.J. Newman, Powerlaw distributions in empirical data. *SIAM Rev.* **51** (4), 661–703 (2009)
- E. Gilbert, K. Karahalios, Predicting tie strength with social media, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI'09*, pp. 211–220 (2009)
- J. Giles, Computational social science: making the links. *Nature* **488**, 448–450 (2012)
- M.S. Granovetter, The strength of weak ties. *Amer. J. Sociol.* **78**(6), 1360–1380 (1973)
- M. Granovetter, The strength of weak ties: a network theory revisited. *Sociol. Theory* **1**(1983), 201–233 (1983)
- J.H.B. de Groot, M.A.M. Smeets, A. Kaldewaij, M.J.A. Duijndam, G.R. Semin, Chemosignals communicate human emotions. *Psychol. Sci.* **23**(11), 1417–1424 (2012). doi:[10.1177/0956797612445317](https://doi.org/10.1177/0956797612445317)
- M.J. Hove, J.L. Risen, It's all in the timing: interpersonal synchrony increases affiliation. *Soc. Cognit.* **27**, 949–960. doi:[10.1521/soco.2009.27.6.949](https://doi.org/10.1521/soco.2009.27.6.949) (2009)
- D. Krackhardt, The strength of strong ties: the importance of philos in organizations, in *Networks and Organizations: Structure, Form, and Action*, ed. by N. Nohria, R. Eccles (Harvard Business School Press, Boston, MA, 1992), pp. 216–239
- P.L. Krapivsky, S. Redner, F. Leyvraz, Connectivity of growing random networks. *Phys. Rev. Lett.* **85**(21), 4629–4632 (2000)
- P.L. Krapivsky, S. Redner, Organization of growing random networks *Phys. Rev. E* **63**, 066123 (14) (2001)
- P.L. Krapivsky, S. Redner, Rate equation approach for growing networks, in *Statistical Mechanics of Complex Networks*, vol. 625, Lecture Notes in Physics, ed. by R. Pastor-Satorras, M. Rubi, A. Diaz-Guilera (2003) pp. 3–22
- P.V. Marsden, K.E. Campbell, Measuring tie strength. *Soc. Forces* **63**(1990), 482–501 (1990)
- M. McPherson, L. Smith-Lovin, J.M. Cook, Birds of a feather: homophily in social networks. *Annu. Rev. Sociol.* **27**, 415–444 (2001)
- A. Mislove, B. Viswanath, K.P. Gummadi, P. Druschel, You are who You know: inferring user profiles in online social networks, in *Proceedings of the WSDM'10*, February 4–6, 2010, New York City, New York, USA, pp. 251–260 (2010)
- M.R. Morris, J. Teevan, K. Panovich, What Do People Ask Their Social Networks, and Why? A Survey Study of Status Message Q&A Behavior. *CHI 2010* (2010)

- M.P. Mundt, L. Mercken, L.I. Zakletskaia, Peer selection and influence effects on adolescent alcohol use: a stochastic actor-based model. *BMC Pediatrics* **12**, 115 (2012)
- M.E.J. Newman, The structure and function of complex networks. *SIAM Rev.* **45**, 167–256 (2003)
- H.V.D. Parunak, E. Downs, A. Yinger, Socially-constrained exogenously-driven opinion dynamics, in *Fifth International, IEEE Conference Self-Adaptive and Self-Organizing Systems (SASO)* (2011)
- E. Pearce, J. Launay, R.I.M. Dunbar, The ice-breaker effect: singing together mediates fast social bonding. *R. Soc. Open Sci.* (2015). doi:[10.1098/rsos.150221](https://doi.org/10.1098/rsos.150221)
- S. Perseguers, M. Lewenstein, A. Acín, J.I. Cirac, Quantum random networks. *Nat. Phys.* **6**, 539–543 (2010)
- D.J. de S. Price, A general theory of bibliometric and other cumulative advantage processes, *J. Amer. Soc. Inform. Sci.* **27**, 292–306 (1976)
- A. Rapoport, Contribution to the theory of random and biased nets. *Bull. Math. Biol.* **19**, 257–277 (1957)
- C.R. Shalizi, A.C. Thomas, Homophily and contagion are generically confounded in observational social network studies. *Sociol. Methods Res.* **40**(2), 211–239 (2011)
- A. Sharpanskykh, J. Treur, Modelling and analysis of social contagion in dynamic networks. *Neurocomput. J.* **146**(2014), 140–150 (2014)
- C.E.G. Steglich, T.A.B. Snijders, M. Pearson, Dynamic networks and behavior: separating selection from influence. *Sociol. Methodol.* **40**, 329–393 (2010)
- T.W. Valente, *Social Networks and Health: Models, Methods, and Applications* (Oxford University Press, New York, 2010)
- T.W. Valente, Network interventions. *Science* **337**(49–53), 2012 (2012)
- W. van Breda, J. Treur, A. van Wissen, Analysis and support of lifestyle via emotions using social media, in *Proceedings of the 4th International Conference on Social Informatics, SocInfo'12*, Lecture Notes in Computer Science (Springer, 2012), pp. 275–291

Part IV
Analysis Methods for Temporal-Causal
Network Models

Chapter 12

Where Is This Going

Verification by Mathematical Analysis

Abstract Within Network-Oriented Modeling based on temporal-causal network models mathematical analysis of the dynamics of the behavior of the network models can be performed. This chapter addresses the analysis of some types of dynamic properties of a temporal-causal network model in an analytical mathematical manner. Properties addressed describe whether some values for the states exist for which no change occurs (equilibria), whether the values for these states converge to such a value as a limit value (attracting equilibria), whether states will show monotonically increasing or decreasing values over time (monotonicity), and whether situations occur in which no convergence takes place but in the end a specific sequence of values is repeated all the time (limit cycle). It is discussed how such analyses can be used for verification of the (implemented) model. Any discrepancies found, suggest there is something wrong in the implementation of the model. In this chapter some methods to analyse such properties of adaptive temporal-causal network models will be described and illustrated for a simple example model, for Hebbian learning, and for adaptive network models for evolving social interaction.

12.1 Introduction

Network-Oriented Modeling based on temporal-causal network models results in dynamical models. Usually emerging dynamic properties of dynamical models can be analysed by conducting simulation experiments. But some specific types of properties can also be found by calculations in a mathematical manner, without performing simulations. Examples of properties that can be analyzed in such a manner are:

- Properties describing whether for some values for the states no change occurs (stationary points or equilibria), and how such values may depend on the values of the parameters of the model and/or the initial values for the states

- Properties describing whether certain states in the model converge to some limit value (equilibria) and how this may depend on the values of the parameters of the model and/or the initial values for the states
- Properties describing whether some state will show monotonically increasing or decreasing values over time (monotonicity)
- Properties describing situations in which no convergence takes place but in the end a specific sequence of values is repeated all the time (limit cycle)

Such properties found in an analytic mathematical manner can be used for verification of the model by checking them for the values observed in simulation experiments. Typically such properties take the form of equations or inequalities for values of one state in relation to values of connected states. If one of these properties is not fulfilled (and the mathematical analysis was done in a correct manner), then there will be some error in the implementation of the model. In some cases, but certainly not always, such equations or inequalities can also be solved in an analytical manner in the sense that the value of one state is expressed in terms of or compared to an arithmetical expression of the parameter values. However, for the purpose of verification solving the equations or inequalities is not required. In this chapter some methods to analyse such properties of models will be described in particular in the setting of adaptive temporal-causal network models. They will be illustrated for the example temporal-causal network model also used in Chap. 2 and for two types of adaptive temporal-causal network models: one based on Hebbian learning, and one based on the homophily principle for dynamic connection strengths in adaptive networks modeling evolving social interaction.

To get the idea, first the general set up is discussed in Sect. 12.2. This is illustrated in Sect. 12.3 by an analysis of a simple example (as discussed in Chap. 2, Sect. 2.4.2), using sum and identity combination functions. In simulations it is observed for this example model that when a constant stimulus level occurs in the world, for each state its activation value increases from 0 to some value that is then kept forever, until the stimulus disappears: an equilibrium state. In subsequent sections three more general examples of this type of analysis for which equilibrium states occur are addressed: for a scaled sum combination function (Sect. 12.4), for Hebbian learning (Sect. 12.5), and for an adaptive network for social interaction based on the homophily principle (Sect. 12.6). In Sect. 12.7 an analysis is discussed for a case in which no equilibrium state occurs, but instead a limit cycle pattern emerges.

12.2 Verifying a Temporal-Causal Network Model by Mathematical Analysis

A stationary point of a state occurs at some point in time if for this time point no change occurs: the graph is horizontal at that point. Stationary points are usually maxima or minima (peaks or dips) but sometimes also other stationary points may

occur. An equilibrium occurs when for all states no change occurs. From the difference or differential equations describing the dynamics for a model it can be analysed when stationary points or equilibria occur. Moreover, it can be found when a certain state is increasing or decreasing when a state is not in a stationary point or equilibrium. First a definition for these notions.

Definition (*stationary point, increase, decrease, and equilibrium*) Let Y be a state

- Y has a *stationary point* at t if $\frac{dY(t)}{dt} = 0$
- Y is *increasing* at t if $\frac{dY(t)}{dt} > 0$
- Y is *decreasing* at t if $\frac{dY(t)}{dt} < 0$

The model is in *equilibrium* at t if every state Y of the model has a stationary point at t . This equilibrium is *attracting* when for any state Y for all values of Y in some neighbourhood of the equilibrium value, they increase when the value is below the equilibrium value and decrease when the value is above the equilibrium value.

To illustrate these notions, recall the example from Chap. 2, with conceptual representation depicted here in Fig. 12.1, and an example simulation shown in Fig. 12.2. Combination functions are the scaled sum function and the identity function, and all connections have weight 1, except the connections to ps_a , which have weight 0.5.

In Fig. 12.2 it can be seen that as a result of the stimulus all states are increasing until time point 35, after which they start to decrease as the stimulus disappears. Just before time point 35 all states are almost stationary. If the stimulus is not taken away after this time point this trend is continued, and an equilibrium state is approximated. The question then is whether these observations based on one or more simulation experiments are in agreement with a mathematical analysis. If it is found out that they are in agreement with the mathematical analysis, then this provides some extent of evidence that the implemented model is correct. If they turn out not to be in agreement with the mathematical analysis, then this indicates that probably there is something wrong, and further inspection and correction has to be initiated.

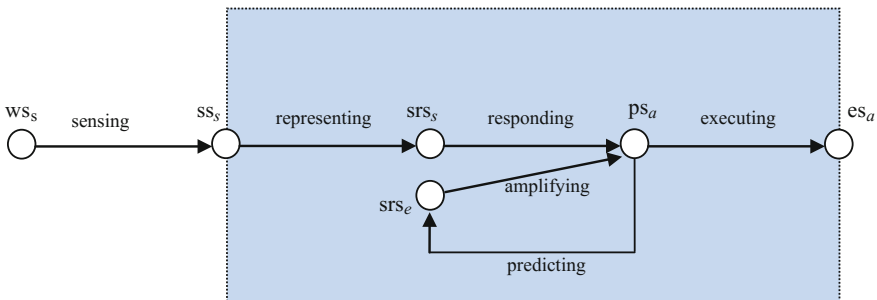


Fig. 12.1 Conceptual representation of an example model

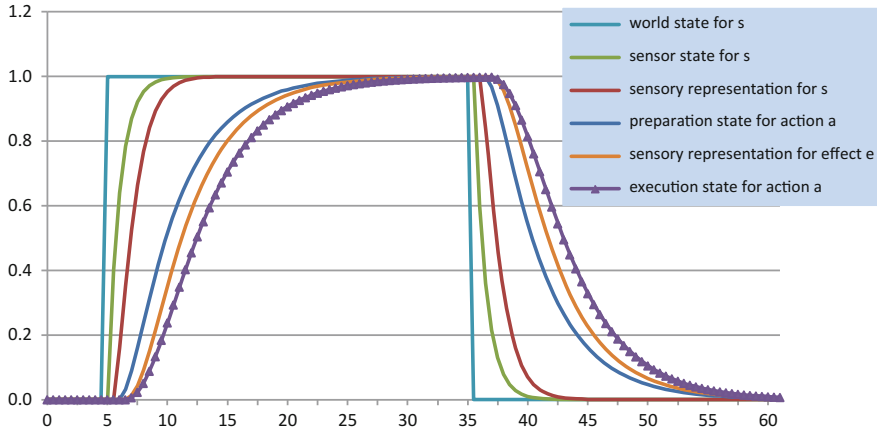


Fig. 12.2 Simulation example for the model depicted in Fig. 12.1 using identity and sum combination functions for all states

Considering the differential equation for a temporal-causal network model more specific criteria can be found:

$$dY(t)/dt = \eta_Y [\text{aggimpact}_Y(t) - Y(t)]$$

with

$$\text{aggimpact}_Y(t) = c_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t))$$

and $X_1, \dots, X_k$ the states with connections to Y . For example, it can be concluded that

$$\begin{aligned} dY(t)/dt > 0 &\Leftrightarrow \eta_Y [\text{aggimpact}_Y(t) - Y(t)] > 0 \Leftrightarrow \\ \text{aggimpact}_Y(t) > Y(t) &\Leftrightarrow c_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) > Y(t) \end{aligned}$$

In this manner the following criteria can be found.

Criteria for a temporal-causal network model: increase, decrease, stationary point and equilibrium

Let Y be a state and $X_1, \dots, X_k$ the states connected toward Y . Then the following hold

Y has a stationary point at $t \Leftrightarrow \text{aggimpact}_Y(t) = Y(t)$

$$\Leftrightarrow c_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) = Y(t)$$

Y is increasing at $t \Leftrightarrow \mathbf{aggimpact}_Y(t) > Y(t)$

$$\Leftrightarrow \mathbf{c}_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) > Y(t)$$

Y is decreasing at $t \Leftrightarrow \mathbf{aggimpact}_Y(t) < Y(t)$

$$\Leftrightarrow \mathbf{c}_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) < Y(t)$$

The model is in equilibrium at $t \Leftrightarrow \mathbf{aggimpact}_Y(t) = Y(t)$ for every state Y

$$\Leftrightarrow \mathbf{c}_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) = Y(t) \text{ for every state } Y$$

These criteria can be used to verify (the implementation of) the model based on inspection of stationary points or equilibria in the following two different manners. Note that in a given simulation the stationary points that are identified are usually approximately stationary; how closely they are approximated depends on different aspects, for example on the step size, or on how long the simulation is done.

Verification by checking the criteria through substitution values from a simulation in the criteria

1. Generate a simulation
2. For a number of states Y identify stationary points with their time points t and state values $Y(t)$
3. For each of these stationary points for a state Y at time t identify the values $X_1(t), \dots, X_k(t)$ at that time of the states $X_1, \dots, X_k$ connected toward Y
4. Substitute all these values $Y(t)$ and $X_1(t), \dots, X_k(t)$ in the criterion $\mathbf{c}_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) = Y(t)$
5. If the equation holds (for example, with an accuracy $< 10^{-2}$), then this test succeeds, otherwise it fails
6. If this test fails, then it has to be explored where the error can be found

This verification method can be illustrated for the example of Figs. 12.1 and 12.2 as follows. For example, consider state ps_a with numerical representation

$$ps_a(t + \Delta t) = ps_a(t) + \eta_{ps_a}[\omega_{\text{responding}}srs_s(t) + \omega_{\text{amplifying}}srs_e(t) - ps_a(t)]\Delta t$$

The equation expressing that a state of ps_a is stationary at time t is

$$\omega_{\text{responding}}srs_s(t) + \omega_{\text{amplifying}}srs_e(t) = ps_a(t)$$

At time point $t = 35$ (where all states are close to stationary) the following values occur: $ps_a(35) = 0.99903$, $srs_s(35) = 1.00000$ and $srs_e(35) = 0.99863$; moreover $\omega_{\text{responding}} = \omega_{\text{amplifying}} = 0.5$. All these values can be substituted in the above equation:

$$0.5 * 1.00000 + 0.5 * 0.99863 = 0.99903$$

$$0.999315 = 0.99903$$

It turns out that the equation is fulfilled with accuracy $< 10^{-3}$. This gives evidence that the model as implemented indeed does what it was meant to do. If this is done for all other states, similar outcomes are found. This gives still more evidence. The step size Δt for the simulation here was 0.5, which is even not so small. For still more accurate results it is advisable to choose a smaller step size. So, having the equations for stationary points for all states provides a means to verify the implemented model in comparison to the model description. The equations for stationary points themselves can easily be obtained from the model description in a systematic manner.

Note that this method works without having to solve the equations, only substitution takes place; therefore it works for any choice of combination function. Moreover, note that the method also works when there is no equilibrium but the values of the states fluctuate all the time, according to a recurring pattern (a limit cycle). In such cases for each state there are maxima (peaks) and minima (dips) which also are stationary. The method can be applied to such a type of stationary points as well; here it is still more important to choose a small step size as each stationary point occurs at just one time point. In Sect. 12.7 it will be discussed how the analysis approach can be applied to such limit cycles.

There is still another method possible that is sometimes proposed; this method is applied for the case of an equilibrium (where all states have a stationary point simultaneously), and is based on solving the equations for the equilibrium values first. This can provide explicit expressions for equilibrium values in terms of the parameters of the model. Such expressions can be used to predict equilibrium values for specific simulations, based on the choice of parameter values. This method provides more than the previous method, but a major drawback is that it cannot be applied in all situations. For example, when logistic combination functions are used it cannot be applied. However, in some cases it still can be useful. The method goes as follows.

Verification by solving the equilibrium equations and comparing predicted equilibrium values to equilibrium values in a simulation

1. Consider the equilibrium equations for all states Y :

$$\mathbf{c}_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) = Y(t)$$

2. Leave the t out and denote the values as constants $\underline{\mathbf{X}}_j$ and $\underline{\mathbf{Y}}$

$$\mathbf{c}_Y(\omega_{X_1,Y}\underline{\mathbf{X}}_1, \dots, \omega_{X_k,Y}\underline{\mathbf{X}}_k) = \underline{\mathbf{Y}}$$

An equilibrium is a solution $\underline{X}_1, \dots, \underline{X}_k$ of the following set of n equilibrium equations in the n states $X_1, \dots, X_n$ of the model:

$$\begin{aligned} c_{X_1}(\omega_{X_1, X_1} \underline{X}_1, \dots, \omega_{X_n, X_1} \underline{X}_n) &= \underline{X}_1 \\ &\dots \\ c_{X_n}(\omega_{X_1, X_n} \underline{X}_1, \dots, \omega_{X_n, X_n} \underline{X}_n) &= \underline{X}_n \end{aligned}$$

3. Solve these equations mathematically in an explicit analytical form: for each state X_i a mathematical formula $\underline{X}_i = \dots$ in terms of the parameters of the model (connection weights and parameters in the combination function $c_{X_i}(\dots)$, such as the scaling factor in a scaled sum combination function); more than one solution is possible
4. Generate a simulation
5. Identify equilibrium values in this simulation
6. If for all states Y the predicted value $\underline{Y}$ from a solution of the equilibrium equations equals the value for Y obtained from the simulation (for example, with an accuracy $< 10^{-2}$), then this test succeeds, otherwise it fails
7. If this test fails, then it has to be explored where the error can be found

In Sect. 12.3 it will be illustrated how this method works for the example depicted in Figs. 12.1 and 12.2. In general, whether or not the equilibrium equations can be solved in an explicit analytical manner strongly depends on the form of the combination functions $c_Y(\dots)$. In a number of specific cases explicit analytical solutions can be found. Three examples of this are addressed in subsequent sections:

- for a (scaled) sum combination function (Sects. 12.3 and 12.4)
- for adaptive network models based on Hebbian learning (Sect. 12.5)
- for adaptive network models based on the homophily principle (Sect. 12.6)

However, there are also many cases in which an explicit analytical solution cannot be determined, for example, when logistic combination functions are used. In such cases equilibria can only be determined either by numerically solving the equations by some numerical approximation method, or by observing the behaviour of the model in simulation experiments. But in the latter case verification is not possible, as then only simulation results are available. An additional drawback is that in such cases specific values for the parameters of the model have to be chosen, whereas in the case of an explicit analytical solution a more generic expression can be obtained which depends, as a function, on the parameter values. For example, for the cases described in Sects. 12.3–12.6 expressions can be found for the equilibrium values in terms of the connection weights (for which no specific values are needed at forehand).

12.3 Mathematical Analysis for Equilibrium States: An Example

Are there cases in which the types of behaviour considered above can be predicted without running a simulation? In particular, can equilibrium values be predicted, and how they depend on the specific values of the parameters of the model (e.g., connection weights, speed factors)? Below, these questions will be answered for a relatively simple example. Indeed it will turn out that in this case it is possible to predict the equilibrium values from the connection weights (the equilibrium values turn out to be independent of the speed factors, as long as these are nonzero).

As a first step, consider the sensor state ss_s .

LP_{ss_s} Sensing a stimulus: determining values for state ss_s

$$dss_s(t)/dt = \eta_{ss_s} [\omega_{\text{sensing}} ws_s(t) - ss_s(t)]$$

Having an equilibrium value means that no change occurs at t : $dss_s(t)/dt = 0$. As it is assumed that both η_{ss_s} is nonzero, this is equivalent to the following equilibrium equation for state ss_s , with $\underline{ws}_s$ and $\underline{ss}_s$ the equilibrium values for the two states ws_s and ss_s .

$$\omega_{\text{sensing}} \underline{ws}_s = \underline{ss}_s$$

In a similar manner this can be done for the other states, resulting in the following equations:

equilibrium of state	equilibrium criterion
ss_s	$\omega_{\text{sensing}} \underline{ws}_s = \underline{ss}_s$
srs_s	$\omega_{\text{representing}} \underline{ss}_s = \underline{srs}_s$
ps_a	$\omega_{\text{responding}} \underline{srs}_s + \omega_{\text{amplifying}} \underline{srs}_e = \underline{ps}_a$
srs_e	$\omega_{\text{predicting}} \underline{ps}_a = \underline{srs}_e$
es_a	$\omega_{\text{executing}} \underline{ps}_a = \underline{es}_a$

These are five equations with 6 unknowns $\underline{ws}_s$, $\underline{ss}_s$, $\underline{srs}_s$, $\underline{ps}_a$, $\underline{srs}_e$, $\underline{es}_a$; however, the state $\underline{ws}_s$ can be considered given as it indicates the external stimulus. So the five equations can be used to find expressions for the equilibrium values for the five other states in terms of the connection weights ω_x and $\underline{ws}_s$. Note that for the sake of simplicity here it is assumed that $\omega_{\text{amplifying}}$ and $\omega_{\text{predicting}}$ are not both 1. Then this can be solved in an explicit analytical manner as follows. First two of them (the first two equations) are expressed in the externally given value $\underline{ws}_s$:

$$\begin{aligned}\underline{\mathbf{ss}}_s &= \omega_{\text{sensing}} \underline{\mathbf{ws}}_s \\ \underline{\mathbf{srs}}_s &= \omega_{\text{representing}} \underline{\mathbf{ss}}_s = \omega_{\text{representing}} \omega_{\text{sensing}} \underline{\mathbf{ws}}_s\end{aligned}$$

Moreover, the third and fourth equation can be solved as follows:

$$\begin{aligned}\omega_{\text{responding}} \underline{\mathbf{srs}}_s + \omega_{\text{amplifying}} \underline{\mathbf{srs}}_e &= \underline{\mathbf{ps}}_a \\ \omega_{\text{predicting}} \underline{\mathbf{ps}}_a &= \underline{\mathbf{srs}}_e\end{aligned}$$

Substitute $\omega_{\text{predicting}} \underline{\mathbf{ps}}_a$ for $\underline{\mathbf{srs}}_e$ in the third equation, resulting in the following equation in $\underline{\mathbf{ps}}_a$ and $\underline{\mathbf{srs}}_s$:

$$\omega_{\text{responding}} \underline{\mathbf{srs}}_s + \omega_{\text{amplifying}} \omega_{\text{predicting}} \underline{\mathbf{ps}}_a = \underline{\mathbf{ps}}_a$$

This can be used to express $\underline{\mathbf{ps}}_a$ in $\underline{\mathbf{srs}}_s$, and subsequently in $\underline{\mathbf{ws}}_s$:

$$\begin{aligned}\omega_{\text{responding}} \underline{\mathbf{srs}}_s &= (1 - \omega_{\text{amplifying}} \omega_{\text{predicting}}) \underline{\mathbf{ps}}_a \\ \underline{\mathbf{ps}}_a &= \omega_{\text{responding}} \underline{\mathbf{srs}}_s / (1 - \omega_{\text{amplifying}} \omega_{\text{predicting}}) \\ &= \omega_{\text{responding}} \omega_{\text{representing}} \omega_{\text{sensing}} \underline{\mathbf{ws}}_s / (1 - \omega_{\text{amplifying}} \omega_{\text{predicting}})\end{aligned}$$

Moreover, by the fourth equation it is found

$$\begin{aligned}\underline{\mathbf{srs}}_e &= \omega_{\text{predicting}} \underline{\mathbf{ps}}_a \\ &= \omega_{\text{predicting}} \omega_{\text{responding}} \omega_{\text{representing}} \omega_{\text{sensing}} \underline{\mathbf{ws}}_s / (1 - \omega_{\text{amplifying}} \omega_{\text{predicting}})\end{aligned}$$

Based on these, the fifth equation can be used to get an expression for $\underline{\mathbf{es}}_a$:

$$\begin{aligned}\underline{\mathbf{es}}_a &= \omega_{\text{executing}} \underline{\mathbf{ps}}_a \\ &= \omega_{\text{executing}} \omega_{\text{responding}} \omega_{\text{representing}} \omega_{\text{sensing}} \underline{\mathbf{ws}}_s / (1 - \omega_{\text{amplifying}} \omega_{\text{predicting}})\end{aligned}$$

Summarizing, all equilibrium values have been expressed in terms of the external state $\underline{\mathbf{ws}}_s$ and the connection weights:

$$\begin{aligned}\underline{\mathbf{ss}}_s &= \omega_{\text{sensing}} \underline{\mathbf{ws}}_s \\ \underline{\mathbf{srs}}_s &= \omega_{\text{representing}} \omega_{\text{sensing}} \underline{\mathbf{ws}}_s \\ \underline{\mathbf{ps}}_a &= \omega_{\text{responding}} \omega_{\text{representing}} \omega_{\text{sensing}} \underline{\mathbf{ws}}_s / (1 - \omega_{\text{amplifying}} \omega_{\text{predicting}}) \\ \underline{\mathbf{srs}}_e &= \omega_{\text{predicting}} \omega_{\text{responding}} \omega_{\text{representing}} \omega_{\text{sensing}} \underline{\mathbf{ws}}_s / (1 - \omega_{\text{amplifying}} \omega_{\text{predicting}}) \\ \underline{\mathbf{es}}_a &= \omega_{\text{executing}} \omega_{\text{responding}} \omega_{\text{representing}} \omega_{\text{sensing}} \underline{\mathbf{ws}}_s / (1 - \omega_{\text{amplifying}} \omega_{\text{predicting}})\end{aligned}$$

For example, if the external stimulus $\underline{ws}_s$ has level 1 this becomes:

$$\underline{ss}_s = \omega_{\text{sensing}}$$

$$\underline{srs}_s = \omega_{\text{representing}} \omega_{\text{sensing}}$$

$$\underline{ps}_a = \omega_{\text{responding}} \omega_{\text{representing}} \omega_{\text{sensing}} / (1 - \omega_{\text{amplifying}} \omega_{\text{predicting}})$$

$$\underline{srs}_e = \omega_{\text{predicting}} \omega_{\text{responding}} \omega_{\text{representing}} \omega_{\text{sensing}} / (1 - \omega_{\text{amplifying}} \omega_{\text{predicting}})$$

$$\underline{es}_a = \omega_{\text{executing}} \omega_{\text{responding}} \omega_{\text{representing}} \omega_{\text{sensing}} / (1 - \omega_{\text{amplifying}} \omega_{\text{predicting}})$$

Moreover if all connection weights are 1, except that $\omega_{\text{responding}} = 0.5$ and $\omega_{\text{amplifying}} = 0.5$, as in the example simulation shown in Chap. 2, Sect. 2.4.2, the values become:

$$\underline{ss}_s = 1$$

$$\underline{srs}_s = 1$$

$$\underline{ps}_a = 0.5/0.5 = 1$$

$$\underline{srs}_e = 0.5/0.5 = 1$$

$$\underline{es}_a = 0.5/0.5 = 1$$

Indeed in the example simulation in Chap. 2, Sect. 2.8, Fig. 2.21 it can be seen that all values go to 1. The solution of the equilibrium equations in terms of the connection weights can be used to predict that when the connection weights have different values, also these equilibrium values will turn out different. Recall that the case $\omega_{\text{amplifying}} = 1$ and $\omega_{\text{predicting}} = 1$ was excluded. In that case the combined third and fourth equation becomes trivial, as $\underline{ps}_a$ is lost from the equation:

$$\omega_{\text{responding}} \underline{srs}_s + \omega_{\text{amplifying}} \omega_{\text{predicting}} \underline{ps}_a = \underline{ps}_a$$

$$\omega_{\text{responding}} \underline{srs}_s + \underline{ps}_a = \underline{ps}_a$$

$$\omega_{\text{responding}} \underline{srs}_s = 0$$

$$\underline{srs}_s = 0$$

Here in the last step it is assumed that $\omega_{\text{responding}} > 0$. As a consequence by the first two equations also $\underline{ss}_s$ and $\underline{ws}_s$ are 0, and by the fourth and fifth equation also the values for the other states. It turns out that in this case there can only be an equilibrium if there is no stimulus at all. As soon as there is a nonzero stimulus in this case that $\omega_{\text{amplifying}} = 1$ and $\omega_{\text{predicting}} = 1$, the values of $\underline{ps}_a$, $\underline{srs}_e$ and $\underline{es}_a$ increase indefinitely to larger and larger values (and in particular do not stay within the interval $[0, 1]$), as can be seen from simulations. Note that there was an additional assumption made that $\omega_{\text{responding}} > 0$. If, in contrast, $\omega_{\text{responding}} = 0$, then still more possibilities for equilibria are available. For example, in that case $\underline{ps}_a$ and $\underline{srs}_e$ can have any value, but they have to be equal due to the fourth equation, but this value is independent of the values of $\underline{ws}_s$, $\underline{ss}_s$ and $\underline{srs}_s$, as there is no nonzero connection between these parts of the graph. So, this would not be a very relevant case.

The analysis above can also be done to find out whether or not the activation level of a state is increasing. As a first step, again consider the sensor state ss_s .

LP_{ss_s} Sensing a stimulus: determining values for state ss_s

$$\begin{aligned} dss_s(t)/dt &= \eta_{ss_s} [\omega_{\text{sensing}} ws_s(t) - ss_s(t)] \\ ss_s(t + \Delta t) &= ss_s(t) + \eta_{ss_s} [\omega_{\text{sensing}} ws_s(t) - ss_s(t)] \Delta t \end{aligned}$$

The activation value increases means

$$dss_s(t)/dt > 0 \quad \text{or} \quad ss_s(t + \Delta t) > ss_s(t)$$

This is equivalent to:

$$\omega_{\text{sensing}} ws_s(t) - ss_s(t) > 0$$

This in turn is equivalent to the criterion that the impact on ss_s is higher than the current activation value:

$$\omega_{\text{sensing}} ws_s(t) > ss_s(t)$$

For example, when $ws_s(t) = 1$ and $\omega_{\text{sensing}} = 1$, then the criterion $\omega_{\text{sensing}} ws_s(t) > ss_s(t)$ indicates the activation of state ss_s will increase as long as it did not reach the value 1 yet. This gives as additional information that the equilibrium value 1 of sensor state ss_s is *attracting*: the value goes in that direction as long as it was not reached.

In a similar manner this can be done for the other states, thus obtaining the following criteria:

state	is increasing if and only if
ss_s	$\omega_{\text{sensing}} ws_s(t) > ss_s(t)$
srs_s	$\omega_{\text{representing}} ss_s(t) > srs_s(t)$
ps_a	$\omega_{\text{responding}} srs_s(t) + \omega_{\text{amplifying}} srs_e(t) > ps_a(t)$
srs_e	$\omega_{\text{predicting}} ps_a(t) > srs_e(t)$
es_a	$\omega_{\text{executing}} ps_a(t) > es_a(t)$

12.4 Mathematical Analysis for Equilibrium States: Scaled Sum Combination Function

The approach described above can be applied easily for the case of a scaled sum combination function $c_i(\dots)$ for each state X_i ; such a scaled sum function $\text{ssum}_{\lambda_i}(\dots)$ with scaling factor λ_i is defined as

$$\mathbf{ssum}_{\lambda_i}(V_1, \dots, V_k) = (V_1 + \dots + V_k)/\lambda_i$$

Suppose the differential equation for some state X_i connected to states X_j is given by

$$dX_i/dt = \eta_i[\mathbf{aggimpact}_i(X_1, \dots, X_k) - X_i]$$

where

$$\begin{aligned} \mathbf{aggimpact}_i(X_1, \dots, X_k) &= c_i(\omega_{1,i}X_1, \dots, \omega_{k,i}X_k) = \mathbf{ssum}_{\lambda_i}(\omega_{1,i}X_1, \dots, \omega_{k,i}X_k) \\ &= (\omega_{1,i}X_1 + \dots + \omega_{k,i}X_k)/\lambda_i \end{aligned}$$

with $\omega_{j,i}$ the specific weights for the connections from X_j to X_i . In this case the following holds:

$$\begin{array}{lll} \text{Increasing } X_i : & X_i(t + \Delta t) > X_i(t) & \Leftrightarrow (\omega_{1,i}X_1(t) + \dots + \omega_{k,i}X_k(t))/\lambda_i > X_i(t) \\ \text{Equilibrium of } X_i : & X_i(t + \Delta t) = X_i(t) & \Leftrightarrow (\omega_{1,i}X_1(t) + \dots + \omega_{k,i}X_k(t))/\lambda_i = X_i(t) \\ \text{Decreasing } X_i : & X_i(t + \Delta t) < X_i(t) & \Leftrightarrow (\omega_{1,i}X_1(t) + \dots + \omega_{k,i}X_k(t))/\lambda_i < X_i(t) \end{array}$$

In particular, the equilibrium equations for the states X_i are

$$\begin{aligned} (\omega_{1,i}\mathbf{X}_1 + \dots + \omega_{k,i}\mathbf{X}_k)/\lambda_i &= \mathbf{X}_i \\ &\dots \\ (\omega_{1,k}\mathbf{X}_1 + \dots + \omega_{k,k}\mathbf{X}_k)/\lambda_k &= \mathbf{X}_k \end{aligned}$$

This means that in an equilibrium state the value $\mathbf{X}_i$ for a state X_i may be a weighted average of the equilibrium values $\mathbf{X}_j$ for the states X_j , in particular when

$$\lambda_i = \omega_{1,i} + \dots + \omega_{k,i}$$

Note that always at least one solution exists: when all are 0. But it is usually more interesting to know whether nonzero solutions exist.

The equilibrium equations are equivalent to

$$\begin{aligned} \omega_{1,1}\mathbf{X}_1 + \dots + \omega_{k,1}\mathbf{X}_k &= \lambda_1\mathbf{X}_1 \\ &\dots \\ \omega_{1,k}\mathbf{X}_1 + \dots + \omega_{k,k}\mathbf{X}_k &= \lambda_k\mathbf{X}_k \end{aligned}$$

or

$$\begin{aligned}
 (\omega_{1,1} - \lambda_1)\underline{\mathbf{X}}_1 + \omega_{2,1}\underline{\mathbf{X}}_2 + \cdots + \omega_{k,1}\underline{\mathbf{X}}_k &= 0 \\
 \cdots \\
 \omega_{1,i}\underline{\mathbf{X}}_1 + \cdots + \omega_{i-1,i}\underline{\mathbf{X}}_{i-1} + (\omega_{i,i} - \lambda_i)\underline{\mathbf{X}}_i + \omega_{i+1,i}\underline{\mathbf{X}}_{i+1} + \cdots + \omega_{k,i}\underline{\mathbf{X}}_k &= 0 \\
 \cdots \\
 \omega_{1,k}\underline{\mathbf{X}}_1 + \cdots + \omega_{k-1,k}\underline{\mathbf{X}}_{k-1} + (\omega_{k,k} - \lambda_k)\underline{\mathbf{X}}_k &= 0
 \end{aligned}$$

In general these linear equilibrium equations can be solved analytically, which in principle can provide symbolic expressions for the equilibrium values of X_j in terms of the connection weights $\omega_{j,i}$ and the scaling factor λ_i . However, for more than two states ($k > 2$) such expressions may tend to become more and more complex, but this depends on the number of these $\omega_{j,i}$ which are nonzero, i.e., how many connections between the states exist. For example, if all states have only one incoming and one outgoing connection (a cascade or loop), then these equations can easily be solved. In some cases no nonzero solution exists. This happens, for example, when the values of the parameters are such that two of the equations in a sense contradict each other, as in the equations $X_1 - 2X_2 = 0$ and $X_1 - 3X_2 = 0$.

In some cases some properties of equilibrium values can be derived. For well-connected temporal-causal network models based on scaled sum functions with as scaling factor the sum of the weights of the incoming connections it can be derived that all states have the same equilibrium value.

Definition 1 A network is called *strongly connected* if for every two nodes A and B there is a directed path from A to B and vice versa.

Lemma 1 Let a temporal-causal network model be given based on scaled sum functions:

$$\mathbf{d}Y/\mathbf{d}t = \eta_Y \left[\sum_{X, \omega_{X,Y} > 0} \omega_{X,Y} X / \sum_{X, \omega_{X,Y} > 0} \omega_{X,Y} - Y \right]$$

Then the following hold.

- (a) If for some state Y at time t for all states X connected toward Y it holds $X(t) \geq Y(t)$, then $Y(t)$ is increasing at t : $\mathbf{d}Y(t)/\mathbf{d}t \geq 0$; if for all states X connected toward Y it holds $X(t) \leq Y(t)$, then $Y(t)$ is decreasing at t : $\mathbf{d}Y(t)/\mathbf{d}t \leq 0$.
- (b) If for some state Y at time t for all states X connected toward Y it holds $X(t) \geq Y(t)$, and at least one state X connected toward Y exists with $X(t) > Y(t)$ then $Y(t)$ is strictly increasing at t : $\mathbf{d}Y(t)/\mathbf{d}t > 0$. If for some state Y at time t for all states X connected toward Y it holds $X(t) \leq Y(t)$, and at least one state X connected toward Y exists with $X(t) < Y(t)$ then $Y(t)$ is strictly decreasing at t : $\mathbf{d}Y(t)/\mathbf{d}t < 0$.

Proof of Lemma 1 (a) From the differential equation for $Y(t)$

$$\begin{aligned} dY/dt &= \eta_Y [\sum_{X, \omega_{X,Y} > 0} \omega_{X,Y} X / \sum_{X, \omega_{X,Y} > 0} \omega_{X,Y} - Y] \\ &= \eta_Y [\sum_{X, \omega_{X,Y} > 0} \omega_{X,Y} X - \sum_{X, \omega_{X,Y} > 0} \omega_{X,Y} Y] / \sum_{X, \omega_{X,Y} > 0} \omega_{X,Y} \\ &= \eta_Y [\sum_{X, \omega_{X,Y} > 0} \omega_{X,Y} (X - Y)] / \sum_{X, \omega_{X,Y} > 0} \omega_{X,Y} \end{aligned}$$

it follows that $dY(t)/dt \geq 0$, so $Y(t)$ is increasing at t . Similar for decreasing.

(b) In this case it follows that $dY(t)/dt > 0$, so $Y(t)$ is strictly increasing. Similar for decreasing. $\square$

Theorem 1 (convergence to one value) *Let a strongly connected temporal-causal network model be given based on scaled sum functions:*

$$dY/dt = \eta_Y [\sum_{X, \omega_{X,Y} > 0} \omega_{X,Y} X / \sum_{X, \omega_{X,Y} > 0} \omega_{X,Y} - Y]$$

Then for all states X and Y the equilibrium values $\underline{X}$ and $\underline{Y}$ are equal: $\underline{X} = \underline{Y}$. Moreover, this equilibrium state is attracting.

Proof of Theorem 1 Take a state Y with highest value $\underline{Y}$. Then for all states X it holds $\underline{X} \leq \underline{Y}$. Suppose for some state X connected toward Y it holds $\underline{X} < \underline{Y}$. Take a time point t and assume $Z(t) = \underline{Z}$ for all states Z . Now apply Lemma 1b) to state Y . It follows that $dY(t)/dt < 0$, so $Y(t)$ is not in equilibrium for this value $\underline{Y}$. This contradicts that this $\underline{Y}$ is an equilibrium value for state Y . Therefore the assumption that for some state X connected toward Y it holds $\underline{X} < \underline{Y}$ cannot be true. This shows that $\underline{X} = \underline{Y}$ for all states connected toward Y . Now this argument can be repeated for all states connected toward Y instead of X . By iteration every other state in the network is reached, due to the strong connectivity assumption; it follows that all other states X in the temporal causal network model have the same equilibrium value $\underline{X}$ as $\underline{Y}$. From Lemma 1b) it follows that such an equilibrium state is attracting: if for any state the value is deviating it will move to the equilibrium value. $\square$

12.5 Mathematical Analysis for Equilibrium States: Hebbian Learning

It can also be analysed from the difference or differential equation when a Hebbian adaptation process has an equilibrium and when it increases or decreases. More specifically, assume the following dynamic model for Hebbian learning (e.g., Hebb 1949) for the strength ω of a connection from a state X_1 to a state X_2 with maximal connection strength 1, learning rate $\eta > 0$, and extinction rate $\zeta \geq 0$ (here $X_1(t)$ and

$X_2(t)$ denote the activation levels of the states X_1 and X_2 at time t ; sometimes the t is left out of $X_i(t)$ and simply X_i is written). Also see Chap. 2, Sect. 2.10.

$$\begin{aligned}\omega(t + \Delta t) &= \omega(t) + [\eta X_1(t)X_2(t)(1 - \omega(t)) - \zeta\omega(t)]\Delta t \\ \mathbf{d}\omega(t)/\mathbf{d}t &= \eta X_1X_2(1 - \omega(t)) - \zeta\omega(t)\end{aligned}$$

Note that also for the states X_1 and X_2 equations may be given, but here the focus is on ω . From the expressions for ω it can be analysed when each of the following cases occurs:

$$\begin{array}{lll}\text{Increasing } \omega : & \mathbf{d}\omega(t)/\mathbf{d}t > 0 & \Leftrightarrow \eta X_1X_2(1 - \omega(t)) - \zeta\omega(t) > 0 \\ \text{Stationary point of } \omega : & \mathbf{d}\omega(t)/\mathbf{d}t = 0 & \Leftrightarrow \eta X_1X_2(1 - \omega(t)) - \zeta\omega(t) = 0 \\ \text{Decreasing } \omega : & \mathbf{d}\omega(t)/\mathbf{d}t < 0 & \Leftrightarrow \eta X_1X_2(1 - \omega(t)) - \zeta\omega(t) < 0\end{array}$$

12.5.1 Analysis of Increase, Decrease or Equilibrium for Hebbian Learning Without Extinction

To keep things a bit simple for a first analysis, for the special case that there is no extinction ($\zeta = 0$), this easily leads to the following criteria

Increasing ω :

$$\eta X_1X_2(1 - \omega(t)) > 0 \quad \Leftrightarrow \quad \omega(t) < 1 \text{ and both } X_1 > 0 \text{ and } X_2 > 0$$

Stationary point of ω :

$$\eta X_1X_2(1 - \omega(t)) = 0 \quad \Leftrightarrow \quad \omega(t) = 1 \quad \text{or } X_1 = 0 \text{ or } X_2 = 0$$

Decreasing ω :

$$\eta X_1X_2(1 - \omega(t)) < 0 \text{ this is never the case, as always } X_i \geq 0 \text{ and } \omega(t) \leq 1$$

So, in case that there is no extinction, the only stationary point is when $\omega = 1$, and as long as this value was not reached yet and both $X_1 > 0$ and $X_2 > 0$, the value of ω increases: the stationary point is attracting. Note that when $X_1 = 0$ or $X_2 = 0$, also a stationary point for ω can be found: no (further) learning takes place; the value of ω stays the same independent of which value it has, so in this case any value is a stationary point value. In simulations this indeed can be observed: as long as both $X_1 > 0$ and $X_2 > 0$ the value of ω keeps on increasing until it reaches 1, but if $X_1 = 0$ or $X_2 = 0$ then ω always stays the same.

12.5.2 Analysis of Increase, Decrease or Equilibrium for Hebbian Learning with Extinction

As a next step this analysis is extended to the case with extinction $\zeta > 0$. In this case the analysis requires slightly more work; here for convenience the t is left out of the expressions.

Increasing ω :

$$\begin{aligned}
 & \eta X_1 X_2 (1 - \omega) - \zeta \omega \\
 \Leftrightarrow & \eta X_1 X_2 - \eta X_1 X_2 \omega - \zeta \omega > 0 \\
 \Leftrightarrow & \eta X_1 X_2 - (\zeta + \eta X_1 X_2) \omega > 0 \\
 \Leftrightarrow & (\zeta + \eta X_1 X_2) \omega < \eta X_1 X_2 \\
 \Leftrightarrow & \omega < \frac{\eta X_1 X_2}{\zeta + \eta X_1 X_2} \\
 \Leftrightarrow & \omega < \frac{1}{1 + \zeta / (\eta X_1 X_2)} \quad (\text{when both } X_1 > 0 \text{ and } X_2 > 0)
 \end{aligned}$$

Note that when $X_1 = 0$ or $X_2 = 0$, the value of ω is never increasing. Similarly the following criteria can be found.

Stationary point of ω :

$$\begin{aligned}
 \eta X_1 X_2 (1 - \omega) - \zeta \omega = 0 & \Leftrightarrow \omega = \frac{\eta X_1 X_2}{\zeta + \eta X_1 X_2} \\
 \Leftrightarrow \omega = \frac{1}{1 + \zeta / (\eta X_1 X_2)} & \quad (\text{when both } X_1 > 0 \text{ and } X_2 > 0) \\
 \eta X_1 X_2 (1 - \omega) - \zeta \omega = 0 & \Leftrightarrow \omega = 0 \quad (\text{when } X_1 = 0 \text{ or } X_2 = 0, \text{ and } \zeta > 0)
 \end{aligned}$$

Decreasing ω :

$$\begin{aligned}
 \eta X_1 X_2 (1 - \omega) - \zeta \omega < 0 & \Leftrightarrow \omega > \frac{\eta X_1 X_2}{\zeta + \eta X_1 X_2} \\
 \Leftrightarrow \omega > \frac{1}{1 + \zeta / (\eta X_1 X_2)} & \quad (\text{when both } X_1 > 0 \text{ and } X_2 > 0) \\
 \eta X_1 X_2 (1 - \omega) - \zeta \omega < 0 & \Leftrightarrow \text{always} \quad (\text{when } X_1 = 0 \text{ or } X_2 = 0, \text{ and } \zeta > 0, \omega > 0)
 \end{aligned}$$

In this more general case with extinction, depending on the values of X_1 and X_2 there may be a positive stationary point value (when both $X_1 > 0$ and $X_2 > 0$) but when $\zeta > 0$ this value is < 1 . Also 0 is a stationary point value (when $X_1 = 0$ or $X_2 = 0$). This looks similar to the case without extinction. Moreover, as before, the value of ω increases when it is under the positive stationary point value and it decreases when it is above this value (it is attracting); for example patterns, see Figs. 12.3 and 12.4.

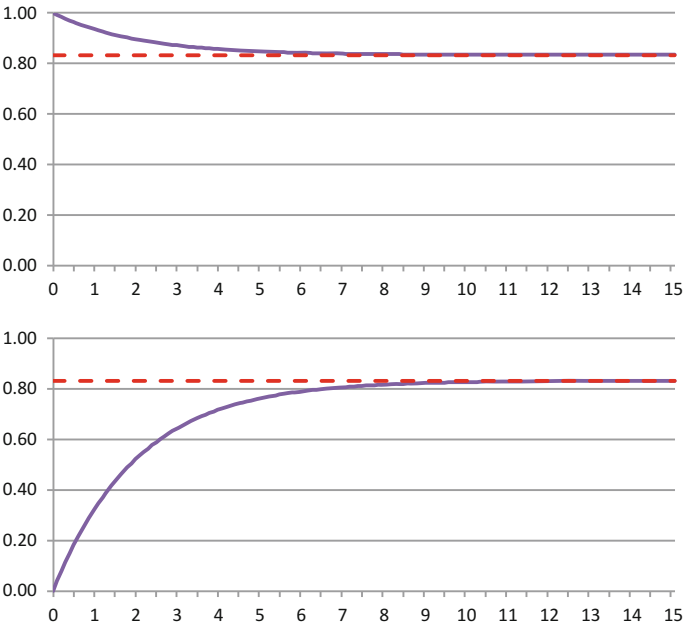


Fig. 12.3 Hebbian learning for $\eta = 0.4$, $\zeta = 0.08$, $\Delta t = 0.1$, and activation levels $X_1 = 1$ and $X_2 = 1$. Equilibrium value 0.83 (*dotted line*)

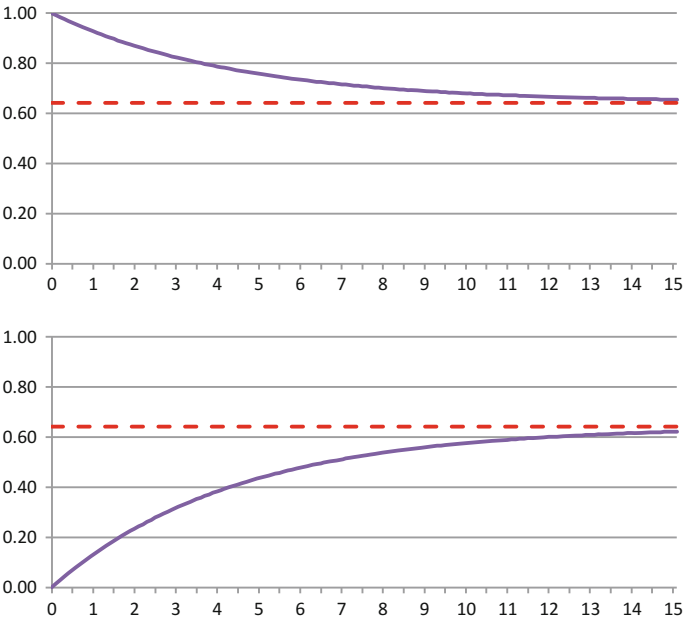


Fig. 12.4 Hebbian learning for $\eta = 0.4$, $\zeta = 0.08$, $\Delta t = 0.1$, and activation levels $X_1 = 0.6$ and $X_2 = 0.6$. Equilibrium value 0.64 (*dotted line*)

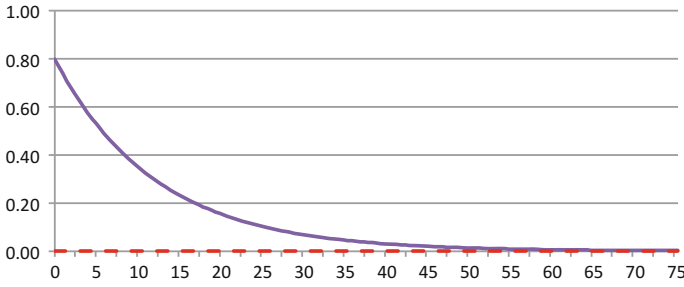


Fig. 12.5 Pure extinction for $\eta = 0.4$, $\zeta = 0.08$, $\Delta t = 0.1$, and activation levels $X_1 = X_2 = 0$; equilibrium value 0

Note that this time this positive stationary point value (indicated by the dotted line) is lower than 1. It may be close to 1, but when $\zeta > 0$ it never will be equal to 1. In fact the maximal value of this stationary point is when both $X_1 = 1$ and $X_2 = 1$, in which case the value is

$$\frac{1}{1 + \zeta/\eta}$$

For example, for $\eta = 0.4$, $\zeta = 0.02$, and $X_1 = 1$ and $X_2 = 1$, the positive stationary point value for ω is about 0.95. Another example is $\eta = 0.4$, $\zeta = 0.08$, and $X_1 = 1$ and $X_2 = 1$, in which case the stationary point value is 0.83. The graphs in Fig. 12.2 show what happens below this stationary point and above it. If for the same settings for η and ζ , the activation levels are lower ($X_1 = 0.6$ and $X_2 = 0.6$), then the stationary point value is lower too (0.64), and the learning is much slower, as is shown in Fig. 12.3.

So, it is found that the positive stationary point value occurs for $X_1 > 0$ and $X_2 > 0$, and in that case this stationary point is attracting. In contrast, the stationary point value 0 does not occur for $X_1 > 0$ and $X_2 > 0$, but it does occur for $X_1 = 0$ or $X_2 = 0$, in which case no positive stationary point value occurs. In this case pure extinction occurs: ω is attracted by the stationary point value 0; this pattern is different from the case without extinction. For an example of such a pure extinction process, see Fig. 12.5. Note that, given the lower value of the extinction rate ζ , the extinction process takes a much longer time than the learning process.

12.5.3 How Much Activation Is Needed to Let ω Increase?

From a different angle, another question that can be addressed is for a given value of ω , how high the value $X_1 X_2$ should be in order to let ω become higher. This can be determined in a similar manner as follows:

Increasing ω : $\omega < \frac{1}{1 + \zeta/\eta(X_1 X_2)} \Leftrightarrow (1 + \zeta/\eta X_1 X_2)\omega < 1$

$$\Leftrightarrow 1 + \zeta/(\eta X_1 X_2) < 1/\omega$$

$$\Leftrightarrow \zeta/(\eta X_1 X_2) < 1/\omega - 1 = (1 - \omega)/\omega$$

$$\Leftrightarrow 1/(X_1 X_2) < \frac{\eta}{\zeta}(1 - \omega)/\omega$$

$$\Leftrightarrow X_1 X_2 > \frac{\zeta}{\eta}\omega/(1 - \omega)$$

So, for activation levels X_1 and X_2 with $X_1 X_2 > \frac{\zeta}{\eta}\omega/(1 - \omega)$, further learning takes place, and below this value extinction dominates and will decrease the level of ω .

12.6 Mathematical Analysis for Equilibrium States: Homophily Principle

The connections between persons in a network describing social interaction may change over time based on the *homophily principle*: the closer the states of the interacting persons, the stronger the connections of the persons will become. In Chap. 11, Sect. 11.7 it is shown how this principle may be formalized with as a general template

$$\mathbf{d}\omega_{A,B}/\mathbf{d}t = \eta_{A,B}[\mathbf{c}_{A,B}(X_A, X_B, \omega_{A,B}) - \omega_{A,B}]$$

for some combination function $\mathbf{c}_{A,B}(V_1, V_2, W)$ for which it is assumed that $\mathbf{c}_{A,B}(V_1, V_2, 0) \geq 0$ and $\mathbf{c}_{A,B}(V_1, V_2, 1) \leq 1$.

The example used in this section is

$$\mathbf{c}_{A,B}(V_1, V_2, W) = W + (\tau_{A,B}^2 - (V_1 - V_2)^2)W(1 - W)$$

In this case

$$\mathbf{d}\omega_{A,B}/\mathbf{d}t = \eta_{A,B}(\tau_{A,B}^2 - (X_A - X_B)^2)\omega_{A,B}(1 - \omega_{A,B})$$

In this section it is analysed which equilibrium values $\underline{\omega}_{A,B}$ can occur for $\omega_{A,B}(t)$ and when $\omega_{A,B}(t)$ is increasing or decreasing.

The standard approach is to derive an inequality or equation from the differential equation by putting $\mathbf{d}\omega_{A,B}(t)/\mathbf{d}t = 0$, $\mathbf{d}\omega_{A,B}(t)/\mathbf{d}t \geq 0$ or $\mathbf{d}\omega_{A,B}(t)/\mathbf{d}t \leq 0$. For this case this provides

Increasing $\omega_{A,B}$

$$d\omega_{A,B}(t)/dt \geq 0 \Leftrightarrow \eta_{A,B}(\tau_{A,B}^2 - (X_A - X_B)^2)\omega_{A,B}(1 - \omega_{A,B}) > 0$$

Equilibrium of $\omega_{A,B}$

$$d\omega_{A,B}(t)/dt = 0 \Leftrightarrow \eta_{A,B}(\tau_{A,B}^2 - (X_A - X_B)^2)\omega_{A,B}(1 - \omega_{A,B}) = 0$$

Decreasing $\omega_{A,B}$

$$d\omega_{A,B}(t)/dt \leq 0 \Leftrightarrow \eta_{A,B}(\tau_{A,B}^2 - (X_A - X_B)^2)\omega_{A,B}(1 - \omega_{A,B}) < 0$$

For $\omega_{A,B} = 0$ or $\omega_{A,B} = 1$ the middle condition is fulfilled. This means that $\underline{\omega}_{A,B} = 0$ and $\underline{\omega}_{A,B} = 1$ are equilibrium values. Now assume $0 < \omega_{A,B} < 1$. Then $\omega_{A,B}(1 - \omega_{A,B}) > 0$, and therefore this factor can be left out, and the same applies to $\eta_{A,B} > 0$; this results in:

Increasing $\omega_{A,B}$

$$\tau_{A,B}^2 - (X_A - X_B)^2 > 0 \Leftrightarrow |X_A - X_B| < \tau_{A,B}$$

Equilibrium of $\omega_{A,B}$

$$\tau_{A,B}^2 - (X_A - X_B)^2 = 0 \Leftrightarrow |X_A - X_B| = \tau_{A,B}$$

Decreasing $\omega_{A,B}$

$$\tau_{A,B}^2 - (X_A - X_B)^2 < 0 \Leftrightarrow |X_A - X_B| > \tau_{A,B}$$

This shows that for cases that $|X_A - X_B| < \tau_{A,B}$ the connection keeps on becoming stronger until $\omega_{A,B}$ becomes in equilibrium at 1. Similarly for cases that $|X_A - X_B| > \tau_{A,B}$ the connection keeps on becomes weaker until $\omega_{A,B}$ becomes in equilibrium at 0. This implies that the equilibria $\underline{\omega}_{A,B} = 0$ and $\underline{\omega}_{A,B} = 1$ can both become attracting, but under different circumstances concerning the values of X_A and X_B .

In exceptional situations it could be the case that $|X_A - X_B| = \tau_{A,B}$ in which case $\omega_{A,B}$ is also in equilibrium, with $\omega_{A,B}$ having any value. So in principle the equilibrium equation has three solutions

$$\underline{\omega}_{A,B} = 0 \quad \text{or} \quad \underline{\omega}_{A,B} = 1 \quad \text{or} \quad |X_A - X_B| = \tau_{A,B} \quad \text{and} \quad \underline{\omega}_{A,B} \text{ has any value}$$

In Chap. 11, Sect. 11.7, Fig. 11.14 for such an adaptive network model an example simulation is shown (for a network with nonzero initial weights for the connections between all pairs of persons) where the connection weights all converge to 0 or 1, and during this process clusters are formed of persons with equal levels of their state. For more details, see Sharpanskykh and Treur (2014).

The analysis above can also be done for similar but slightly more complex variants of the model, of which the quadratic variant is described in Sharpanskykh and Treur (2014):

$$\begin{aligned} c_{A,B}(V_1, V_2, W) &= W + \text{Pos}(\eta_{A,B}(\tau_{A,B} - |V_1 - V_2|))(1 - W) - \text{Pos}(-\eta_{A,B}(\tau_{A,B} - |V_1 - V_2|))W \\ c_{A,B}(V_1, V_2, W) &= W + \text{Pos}(\eta_{A,B}(\tau_{A,B}^2 - (V_1 - V_2)^2))(1 - W) - \text{Pos}(-\eta_{A,B}(\tau_{A,B}^2 - (V_1 - V_2)^2))W \\ c_{A,B}(V_1, V_2, W) &= W + \text{Pos}(\eta_{A,B}(0.5 - 1/(1 + e^{-\sigma_{A,B}(|V_1 - V_2| - \tau_{A,B})}))) (1 - W) \\ &\quad - \text{Pos}(-\eta_{A,B}(0.5 - 1/(1 + e^{-\sigma_{A,B}(|V_1 - V_2| - \tau_{A,B})})))W \end{aligned}$$

where $\text{Pos}(x) = (|x| + x)/2$, which returns x when x is positive and 0 when x is negative. As discussed in Chap. 11 these models make that the approaching of the boundaries 0 and 1 of the interval $[0, 1]$ of ω is slow, thus making ω not crossing these boundaries, but ω departing from the neighbourhood of these boundaries is not slow. In Sharpanskykh and Treur (2014) an analysis and example simulations can be found using the second, quadratic model. As part of the analysis, there it is also shown that different equilibrium values $\underline{\mathbf{X}}_A$ and $\underline{\mathbf{X}}_B$ have a distance of at least $\tau_{A,B}$, which implies that at most $1/\tau_{A,B}$ clusters can emerge.

12.7 Mathematical Analysis for Behaviour Ending up in a Limit Cycle Pattern

Sometimes the values of the states of a model do not end up in an equilibrium value, but instead keep on fluctuating all the time, and after some time they do this according to a repeated pattern, called a *limit cycle*. The example model shown in Figs. 12.1 and 12.2 can be extended to show such behaviour; see Fig. 12.6. In this case it is assumed that action a directs the person (e.g., his or her gaze) away from the stimulus s , so that after (full) execution of a stimulus s is not sensed anymore. This type of behaviour can occur as a form of emotion regulation in order to down-regulate a stressful emotion triggered by s (this is sometimes called attentional deployment; e.g., Gross 1998). The effect of this is as follows. The presence of stimulus s leads to high activation levels of sensor state and sensory representation for s , and subsequently for the preparation state and execution state of action a . But then the action leads to its effect in the world which is suppression of the sensor state for s . As a consequence the sensor state and sensory representation for s , and also the preparation state and execution state of action a get low activation levels. The effect is that there is no suppression of sensing the stimulus anymore and therefore all activation levels become high again. And so it goes on and on, forever (see also Fig. 12.8). At a longer time scale this type of pattern may also occur in so-called on-again-off-again relationships. Note that this type of pattern is also an element of the model discussed in Chap. 9. This type of behaviour can be achieved by the following additions to the example model (see Fig. 12.7):

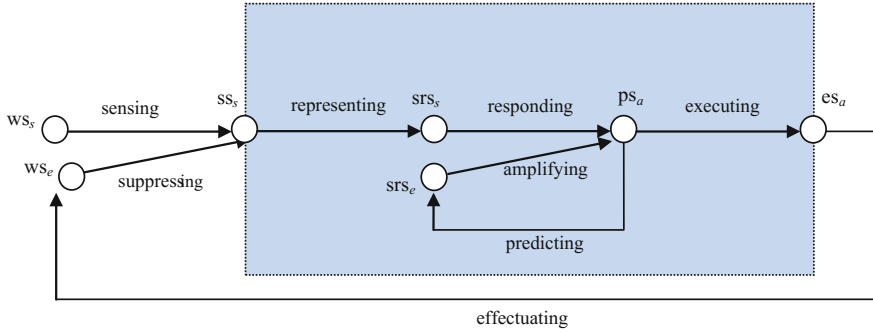


Fig. 12.6 Simple example model incorporating suppression of sensing

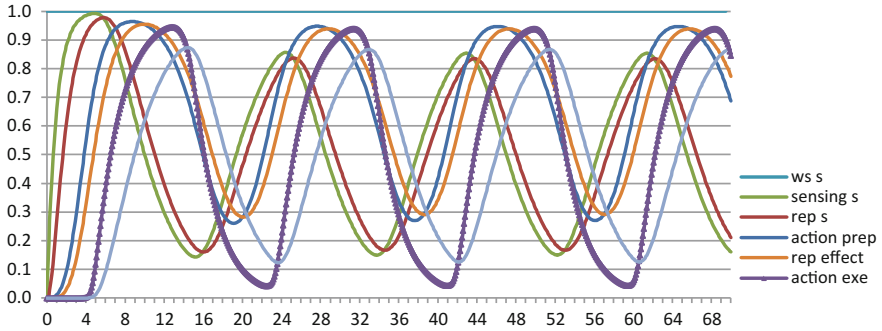


Fig. 12.7 Example simulation showing a limit cycle

- a connection from the execution state es_a of a to the world state ws_e for effect e of action a
- a connection from this world state ws_e for e to the sensor state ss_s of s
- a combination function for the sensor state ss_s of s that models that ws_e makes that s is not sensed

The aggregation used for ss_s is modeled by the following combination function $c_{ss_s}(V_1, V_2)$, where V_1 refers to the impact $\omega_{ws_s, ss_s} ws_s(t)$ from ws_s on ss_s and V_2 to the impact $\omega_{ws_e, ss_s} ws_e(t)$ from ws_e on ss_s :

$$c_{ss_s}(V_1, V_2) = V_1(1 + V_2)$$

Since the connection weight ω_{ws_s, ss_s} is chosen negative (it is a suppressing link), for example -1 , this function makes the sensing of stimulus s inverse proportional to the extent $ws_e(t)$ of avoidance; e.g., sensing s becomes 0 when avoidance e is 1, and V_1 when avoidance e is 0. According to this combination function the difference and differential equation for ss_s are as follows:

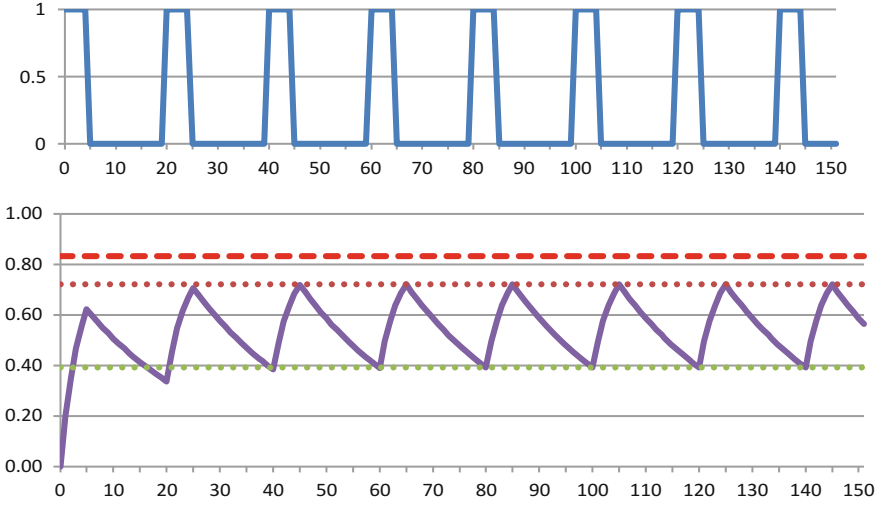


Fig. 12.8 Limit cycle for $d_1 = 5$ (learning), $d_0 = 15$ (pure extinction), and $\eta = 0.2$, $\zeta = 0.04$. Equilibrium value 0.83, $\omega_{\max} = 0.72$, $\omega_{\min} = 0.39$ (dotted lines)

$$\begin{aligned} ss_s(t + \Delta t) &= ss_s(t) + \eta_{ss_s} [\omega_{ws_s, ss_s} ws_s(t)(1 - \omega_{ws_e, ss_s} ws_e(t)) - ss_s(t)] \Delta t \\ dss_s/dt &= \eta_{ss_s} [\omega_{ws_s, ss_s} ws_s(t)(1 - \omega_{ws_e, ss_s} ws_e(t)) - ss_s(t)] \end{aligned}$$

The combination functions for all states with only one connection toward it are the identity function, except for es_a in which case the advanced logistic function **alogistic** $_{\sigma, \tau}(\dots)$ is used. The combination function for ps_a is the sum function or the advanced logistic function **alogistic** $_{\sigma, \tau}(\dots)$.

In Fig. 12.8 an example simulation with the model depicted in Fig. 12.7 clearly shows how a limit cycle pattern emerges, with period 18.5.

Here all connection weights are 1, except the weight of the suppressing connection from ws_e to ss_s , which is -1 . Moreover, the steepness σ and threshold τ for ps_a are 4 and 0.9, respectively, and for es_a they are 40 and 0.7. The step size Δt was 0.1 and the speed factors η for es_a and ws_e were 0.4, and for the other (internal) states η was 1.

For this simulation an analysis of the stationary points has been performed for the maxima and minima in the final stage for all states. Recall from Sect. 12.2 the equation expressing that a state Y is stationary at time t is

$$\mathbf{aggimpact}_Y(t) = Y(t)$$

which is equivalent to

$$\mathbf{c}_Y(\omega_{X_1, Y} X_1(t), \dots, \omega_{X_k, Y} X_k(t)) = Y(t)$$

Table 12.1 Overview of the outcomes of a mathematical analysis for stationary points in a limit cycle

	ws_e	ss_s	srs_s	srs_e	ps_a	es_a
<i>Maxima</i>						
Time point	69.9	61.4	62.2	65.8	64.6	68.3
State value	0.86700	0.85401	0.83455	0.93859	0.94754	0.93975
Aggregated impact	0.86105	0.85001	0.83325	0.93713	0.94703	0.94012
Absolute deviation	0.00595	0.00400	0.00131	0.00146	0.00051	0.00037
<i>Minima</i>						
Time point	60.6	52.2	53.0	57.1	56.1	59.4
State value	0.12553	0.14993	0.16699	0.29153	0.27012	0.04400
Aggregated impact	0.13033	0.15168	0.16689	0.29480	0.27159	0.04317
Absolute deviation	0.00480	0.00175	0.00009	0.00327	0.00147	0.00083

For example, for state ps_a if the combination function chosen is the sum function, the aggregated impact is

$$\text{aggimpact}_Y(t) = \omega_{\text{respondingSRS}_s}(t) + \omega_{\text{amplifyingSRS}_e}(t)$$

Then the stationary point equation expressing that state ps_a is stationary at time t is

$$\omega_{\text{respondingSRS}_s}(t) + \omega_{\text{amplifyingSRS}_e}(t) = ps_a(t)$$

It is such an equation that can be checked for the minima and maxima for each of the states in the final stage of the simulation. Such results are shown in Table 12.1. Here both for the maxima and for the minima the first rows show the time points at which the stationary point occurs. The next row (state value) shows the values of the right hand side of the above equation, followed by rows (aggregated impact) showing the left hand sides of this equation, and then a row with the absolute deviation between the values in the two rows above it.

It turns out that the stationary point equations are fulfilled with an average accuracy over all states and stationary points of 0.002 and a maximal accuracy of 0.006, which both is $<10^{-2}$. This provides evidence that the implemented model is correct in comparison to the model description. In Table 12.1 the more specific numbers are shown for the different states. For the maxima the average deviation is 0.00226, and the maximal absolute deviation is 0.00595 (which occurs for state ws_e). For the minima the average absolute deviation is 0.00204, and the maximal absolute deviation is 0.00480 (which again is for state ws_e). Taken minima and

maxima together, the overall average absolute deviation is 0.00215, and the maximal absolute deviation is 0.00595 (for the maxima of state ws_e).

As another type example of the emergence of limit cycle behaviour, consider that in a realistic context stimuli can be present for some time, but also may be absent for certain periods according to fixed periods, for example, day/night rhythms. As an example, for Hebbian learning, for activations based on stimuli that return from time to time an analysis can be made about when there is enough stimulation over time to achieve or maintain a value for the weight ω of some connection. As an example, see the pattern in Fig. 12.7, where the upper graph shows the levels of both X_1 and X_2 (alternating between 0 and 1) and the lower graph shows how due to these activation periods, the periods of learning ($d_1 = 5$ time units) and pure extinction ($d_0 = 15$ time units) alternate. It turns out that there is a form of convergence not to one specific value of ω , but to a recurring pattern that repeats itself; this is a specific case of a limit cycle, in this case induced by environmental fluctuations. Note that for such cases the verification method described above has problems with the minima or maxima due to the abrupt transition from decrease to increase or conversely.

12.8 Discussion

In this chapter it was discussed how mathematical analysis can be used to find out some properties of the behaviour dynamics of a network model designed according to a Network-Oriented Modeling approach based on temporal-causal networks. The content is based on Treur (2016). An advantage is that it can be used as an additional source of knowledge, independent of a specific implementation of the model. By comparing properties found by mathematical analysis and properties observed in simulation experiments some form of verification can be done. If a discrepancy is found, for example in the sense that the mathematical analysis predicts a certain property but some simulation does not satisfy this property, this can be a reason to inspect the implementation of the model carefully (and/or check whether the mathematical analysis is correct). Having such an option can be fruitful during a development process of a model, as to acquire empirical data for validation of a model may be more difficult or may take a longer time.

The techniques used for such mathematical analysis were adopted from Brauer and Nohel (1969), Lotka (1956), Picard (1891, 1893), Poincaré (1882–1892). In this literature many more techniques can be found than those covered in the current chapter, for example, for the convergence speed for attracting equilibria (e.g., Mathunjwa and Temple 2006), but also for other types of properties. For example, there is underlying theory that proves the existence of certain patterns, for example theorems from Poincaré (1881–1882) that state that under certain circumstances for two-dimensional systems (described by only two differential equations) limit cycles will occur. These are beyond the scope of the current chapter. As some related work

with a different perspective, in Mooij et al. (2013) it is addressed how equilibria of first-order differential equations for the deterministic case can be related to (non-dynamic) Structural Causal Models.

References

- T. Bosse, M. Hoogendoorn, Z.A. Memon, J. Treur, M. Umair, An adaptive model for dynamics of desiring and feeling based on Hebbian learning, in *Proceedings of the Second International Conference on Brain Informatics*, vol. 6334, BF'10. Lecture Notes in Artificial Intelligence, ed. Y. Yao, R. Sun, T. Poggio, J. Liu, N. Zhong, J. Huang (Springer, 2010), pp. 14–28
- F. Brauer, J.A. Nohel, 1969 *Qualitative Theory of Ordinary Differential Equations* (Benjamin, 1969)
- D. Hebb, *The Organisation of Behavior* (Wiley, 1949)
- A.J. Lotka, *Elements of Physical Biology*, 2nd ed. (Williams and Wilkins Co., 1924) (Dover Publications, 1956)
- J.S. Mathunjwa, J. Temple, Convergence behaviour in exogenous growth models. Discussion Paper No. 06/590, Department of Economics, University of Bristol (2006)
- J.M. Mooij, D. Janzing, B. Schölkopf, From differential equations to structural causal models: the deterministic case, in *Proceedings of the 29th Annual Conference on Uncertainty in Artificial Intelligence (UAI-13)*, ed. by in A. Nicholson, P. Smyth (AUAI Press, 2013), pp. 440–448. <http://auai.org/uai2013/prints/papers/24.pdf>
- E. Picard, *Traité d'Analyse*, vol. 1 (1891)
- E. Picard, *Traité d'Analyse*, vol. 2 (1893)
- H. Poincaré, Mémoire sur les courbes définies par une équation différentielle (1881–1882) (On curves defined by differential equations)
- H. Poincaré, *New Methods of Celestial Mechanics*, 3 vols. (1892–1899) English translation (1967)
- A. Sharpanskykh, J. Treur, Modelling and analysis of social contagion in dynamic networks. *Neurocomput. J.* **146**(2014), 140–150 (2014)
- J. Treur, Verification of temporal-causal network models by mathematical analysis. *Vietnam J. Comput. Sci.* (2016). doi:[10.1007/s40595-016-0067-z](https://doi.org/10.1007/s40595-016-0067-z)

Chapter 13

What Is Happening

Identifying and Verifying Emergent Patterns

Abstract In this chapter different types of dynamic properties that can emerge in the behavior dynamics of a temporal-causal network are discussed. It is discussed how they can be expressed both by conceptual and numerical-logical representations, and how they can be checked in a systematic manner for a given behavior trace describing the states of a process over time. Moreover, it is discussed how such checking can be automated by using a dedicated software environment. This can be applied both in the context of monitoring and analysis of a real world process (either off line or on the fly) and in the context of extensive testing, focusing and analysis of a temporal-causal network model and imposed requirements during its development.

13.1 Introduction

When applying Network-Oriented Modeling based on temporal-causal networks such as discussed in Chap. 2, the temporal-causal relations in the network describe the local mechanisms of a process considered from a local perspective. They can be considered as local dynamic properties driving the emerging behaviour dynamics of the overall process. For example, in (Ashby 1960; Port and van Gelder 1995) the emphasis of the numerical representations is on the specification of such local dynamic properties. However, the dynamic patterns as generated in a process over time, can be described in a more global manner, by expressing temporal relations over longer time periods. Such patterns of behaviour dynamics can be considered as emergent phenomena, emerging from the (local) mechanisms that are represented by the causal relations used to describe the model. Descriptions of such dynamic patterns are called *dynamic properties*. As an illustration, some examples of such dynamic properties are:

- eventually action a will be performed
- as long as there is a bad feeling, no action a will be performed
- if action a is performed, then at an earlier point in time stimulus s has been sensed.

Dynamic property expressions are used to describe such patterns; they basically describe a temporal structure: references occur to different time points and order relations between time points, and logical relations such as implications and quantifiers over time are used.

A modeler may work from different viewpoints. One viewpoint is that the local mechanisms are known and incorporated in the model, but the types of patterns that may emerge from them are not known. Then by simulation experiments for different characteristics, as represented by settings for initial values, input from the environment, and/or parameter values, such emergent patterns can be discovered. Subsequently it can be investigated whether such patterns also occur in the real world. If they actually do occur, this contributes to *validation* of the *model*: some evidence has been obtained that the model describes the real world processes well.

Another viewpoint is that at forehand not only the local mechanisms are known and can be incorporated in the model, but also at least some expected global patterns are considered, for which it is assumed that they sometimes occur in the real world. In this case dynamic properties describing such patterns can serve as a kind of requirements for the model. The model will not be considered satisfactory when it cannot generate these patterns for at least some of the characteristics (settings for initial values, input from the environment, and/or parameter values). Such dynamic properties can play a role in a modeling process similar to the role of requirements within a software, knowledge or agent system engineering process, as a way of focussing the process on what the system being developed is expected to provide; e.g., (Pohl 2010; Pohl and Rupp 2011; van Lamsweerde 2009; Herlea et al. 1999, 2005; Ferber et al. 2001).

When the model does generate the patterns of dynamics described by the requirements for certain settings, this can be considered a *verification* of the model *with respect to the requirements*. When these requirements themselves are found to correspond to patterns observed in the real world, this results in *validation* of the *requirements* and when the model in turn was verified and found out to satisfy the requirements, via them validation of the model is obtained.

In both cases a more detailed analysis of dynamic properties describing such patterns is worthwhile, for example to find out for which characteristics exactly they do occur: it is often an interesting challenge to analyse how their emergence relates to local descriptions of underlying mechanisms together with initial values, input from the environment, and their characteristics as expressed by the values of the parameters.

Dynamic properties can be expressed in the form of a conceptual representation or in the form of a numerical-logical representation. A conceptual representation is informal (or semiformal), usually a (structured) natural language expression. A numerical-logical representation is a formal expression. Expressing dynamic properties in a formal format makes it possible to verify whether they hold in some given empirical or simulated scenario in an automated manner. This can be helpful in particular if the properties are complex or if many of such checks have to be done, for example, by analysing the effects of a systematic variation of initial values, environmental input and/or parameter values in a simulation experiment.

In this chapter, first in Sect. 13.2 dynamic properties will be discussed and their relation with temporal-causal network descriptions. Next, in Sect. 13.3 it is discussed how dynamic properties can be related to real world processes, either as a way of validating these properties or as a way of monitoring these real world processes. In Sect. 13.4 it is discussed how dynamic properties can be related to a model descriptions, to obtain verification and/or personalisation of a model. In Sect. 13.5 the way of expressing dynamic properties in the form of a conceptual representation using (structured) natural language is discussed. In Sect. 13.6 it will be discussed how numerical-logical representations for dynamic properties can be expressed. These representations will be illustrated by various types of examples in Sects. 13.7 and 13.8. Finally, in Sect. 13.9 it is discussed how numerical-logical representations of dynamic properties can be checked in an automated manner. Section 13.10 is a brief discussion.

13.2 Dynamic Properties and Temporal-Causal Network Models

In this section it is discussed how dynamic properties can be used to describe processes from an overall perspective, in contrast to temporal-causal network models that describe the underlying local mechanisms of a process. First, in Sect. 13.2.1 the idea of dynamic properties to describe dynamic patterns in the overall dynamics is introduced, and the entailment relation between a model description and dynamic properties is briefly discussed. In Sect. 13.2.2 it is discussed how dynamic properties can be identified and formulated based on emerging dynamic patterns from a given temporal-causal network model, and in Sect. 13.2.3 it is discussed how dynamic properties can play a role as requirements for the design of a model. In Sect. 13.7 different types of dynamic properties are distinguished and illustrated, together with their conceptual and numerical-logical representations.

13.2.1 A Temporal-Causal Network Model Describing Local Dynamics and Dynamic Properties Describing Patterns Emerging in Overall Dynamics

A temporal-causal network model consists of a description of the states, and connections of the network defining the local dynamic of the interactions between states. During simulations using the model an overall process is generated from the local dynamics based on these interactions. In this way a model reflects the way in which overall processes in the real world can be analysed by looking at the underlying local mechanisms. The emerging dynamic patterns shown in such an

overall process cannot be found in the model description itself, but they can be described in a different manner, by dynamic properties of the types as discussed in this chapter. For example, consider a process of gossiping in which local interactions are that each person knowing some rumour tells it to at all other persons he or she knows. A dynamic property of the overall process in a strongly connected network (see also Chap. 12, Sect. 12.4) is that eventually everybody knows the rumour. This dynamic property is not described in the local interaction between two persons, but in some sense results from these local interactions. This provides two different, non-equivalent types of descriptions of the same process.

However, being different descriptions of the same process, the two types of descriptions at least have some relation. This relation can be viewed as an *entailment relation*: the dynamic properties describing a process from an overall perspective are entailed by the description of the temporal-causal network model generating the process, plus initial values for states, environmental input and values for parameters representing specific characteristics. For the gossiping example the fact that every person who knows the rumour will tell it to all other persons who he or she knows, together with an initial state that at least one person knows the rumour, entails the dynamic property that eventually everybody in the network knows the rumour, assuming that the network is strongly connected.

13.2.2 *Identifying Emergent Dynamic Properties for a Given Model*

Although the emerging dynamic patterns shown in an overall process cannot be found in the model description itself, still in some way they are entailed by the local dynamics and some additional information concerning initial values of the states and values of parameters. This entailment relation between model description and emerging patterns is usually not known. Therefore, the emerging patterns are not always predictable. In different simulation scenarios they can be observed, identified, and described in the form of dynamic properties as addressed in this chapter. When such dynamic properties have been described, by systematic simulation experiments it can be explored whether and under which circumstances (initial values, environment, parameter values) they emerge, thus getting some (often only partial) knowledge about the entailment relation between model and overall process.

As an example, one type of dynamic pattern indicates that a process reaches a state in which something has been achieved. Recall the example model from Chap. 2. Recall that this model describes a number of local interactions between states:

- a stimulus s in the world occurs (e.g., a place where they sell ice-cream)
- a stimulus occurring in the world is sensed by the person
- sensing a stimulus connects to a sensory representation of this stimulus
- this sensory representation affects a preparation state for an action a (go there to buy the ice-cream),

- a feeling state is affecting this preparation state as well.
- this feeling state in turn is affected by the preparation state
- when the preparation state has a high activation level, it will make that the action is performed.

Considering these local interactions, there may be reason to identify the following dynamic property for the overall process dynamics:

DP1 Always when stimulus s occurs in the world, at some point in time after this execution of action a will occur

Note that for a state by ‘occur’ here it is meant that the activation level of the state will be above some (high) value, for example, above 0.8.

Is this dynamic property DP1 entailed by the model description? The answer on this question is not straightforward: it is not always the case, it depends. It can be explored by simulation experiments whether and when this dynamic property DP1 occurs, depending on circumstances defined by parameter values used, such as connection weights. For example, if the connection weights from sensory representation of s to the preparation state for a , and from this preparation state to the feeling state are relatively low, the preparation state will never reach a high level, and the action will not occur. This illustrates how the entailment relation usually depends on specific circumstances. Often only systematic simulation experiments are a way to acquire some knowledge about this entailment relation.

13.2.3 Identifying Dynamic Properties Initially as Requirements for a Model

In Sect. 13.2.2 a viewpoint was discussed based on an already given model. In the current section a different viewpoint is addressed: the viewpoint for a model that is still to be designed. Also in such a situation dynamic properties of an overall process can be important. Suppose in the real world it was observed that somebody sees the ice cream and goes there to get it, and the aim is to design a temporal-causal network model that describes this based on internal mechanisms involving a representation of the stimulus and a feeling about getting the ice cream. For this situation dynamic property DP1 described above can be considered a requirement for the model. So, this dynamic property can be identified and expressed at fore-hand, before there is any model. During the design process this requirement can be kept in mind to focus the design process on obtaining a type of model that is aimed for. Here again the entailment relation discussed in Sects. 13.2.1 and 13.2.2 plays a role, but now from a reverse viewpoint: given the required dynamic property how can a model be designed that entails this requirement (under certain circumstances concerning initial values, environmental input and parameter values).

For this case also some intermediate states within the longer lasting process are reached. These can also be expressed by required dynamic properties. Such properties can be used, for example, to analyse the overall process as a number of sub-processes that, each result in the achievement of some milestone. For the above example the sensory representation of s and feeling b could be considered milestones, first from the occurrence of s to sensory representation of s (DP2), then from sensory representation of s to feeling b (DP3), and finally, from feeling b to performing action a (DP4), expressed by the following three dynamic properties:

- DP2** Always, when stimulus s occurs in the world, at some point in time after this a sensory representation of s will occur
- DP3** Always, when a sensory representation of s occurs, at some point in time after this feeling b will occur
- DP4** Always, when feeling b occurs, at some point in time after this execution of action a will occur

These more specific or refined requirements can provide still more focus for the design process: try to design the model so that the refined requirements DP2, DP3 and DP4 are fulfilled. This suggests to focus on designing specific parts of the model that realize a way to obtain a sensory representation of stimulus s , a way to generate feeling b , and a way to initiate action a after this feeling, respectively.

Note that properties of the type discussed above are usually not the only relevant type of requirements for a model. For example, satisfying the following dynamic properties may be considered as required as well (for example, to prevent the occurrence of the undesirable situation that the action is performed for no good reason):

- DP5** Always, when execution of action a occurs, at some point in time before this, feeling b occurred
- DP6** Always, when a feeling b occurs, at some point in time before this, the sensory representation of s occurred
- DP7** The feeling level for b is always at most 1
- DP8** When always the stimulus s in the world has level 1, for any two time points the level for feeling b at the earlier time point is at most the level for feeling b at the later time point (monotonically increasing feeling level)

For another example of the use of requirements or constraints in the design and fine-tuning of a temporal-causal network model, see (Thilakaratne 2015).

13.3 Dynamic Properties Versus Real World Dynamics: Validation, Monitoring, and Analysis

Expressing dynamic properties for an overall process can be based on different sources. Not all sources lead to realistic dynamic properties, in the sense that in the real world such properties hold. For example, one possible source is imagination.

Using this as a source may lead to adequate dynamic properties, but it may equally well lead to dynamic properties that may be interesting for artificially engineered processes, but do not apply to real world processes. If real world processes are meant to be modeled, for example, human mental processes, or social processes, more care is needed to achieve some indication or guarantee that the dynamic properties are realistic. This issue concerns *validation* of dynamic properties and is discussed in Sects. 13.3.1 and 13.3.2.

Often realistic dynamic properties only occur in the real world under certain circumstances. Sometimes it is useful to detect when they occur, for example, because such circumstances need some specific interventions. In such cases dynamic properties can be used for *monitoring* real world processes. Such usages of dynamic properties are discussed in Sect. 13.3.3.

A software environment is available with an editor to express dynamic properties and with a checker to automatically check such dynamic properties against given traces (or sets of traces) describing a particular process over time by representing what happens exactly at which points in time. Such an environment can be used for different types of applications; in particular, it can be used in the types of work described in Sects. 13.3.1 and 13.3.3 to automate the checking. The TTL environment is such an environment; it is briefly described in Sect. 13.9.

13.3.1 *Validating Dynamic Properties Against Actual Real World Processes*

Different methods can be used to validate dynamic properties against information about real world processes. When *primary sources* of information are available, in the form of recorded *empirical traces* over time representing what happens exactly at which points in time, such validation may take place in an automated manner; this is addressed in the current section. In other cases, there only may be *secondary sources*, for example publications in empirically based literature in which patterns over time are reported. This will be addressed in Sect. 13.3.2.

Suppose information on real world processes is available, in the form of recorded empirical traces over time representing what happens exactly at which points in time. These traces can be represented in a formal manner, for example for a given trace tr by describing a relation $\mathbf{trace}(tr, X, t, V)$ between states X , time points t and values V . These formal data about the trace and a considered dynamic property can be used as input for a software environment that automatically checks whether the dynamic property holds for the trace; see Sect. 13.9.

13.3.2 Validating Dynamic Properties Against Patterns Reported in Literature

In many cases, there only may be secondary sources of information about real world processes, for example, in the form of publications in empirical literature in which patterns over time are reported, but without giving exact time points and time durations. Or, in such literature theories or hypotheses are described which themselves have a grounding in empirical research. In such cases of secondary sources still validation is possible, but more linguistically and not automatically supported: the dynamic property is linguistically compared to what is described in this literature. This literature can also be used to find inspiration to express the dynamic property.

13.3.3 Monitoring and Analysis of Real World Processes Using Dynamic Properties

When certain dynamic properties are realistic, but only occur in the real world under certain circumstances, some other application can be made. To detect when they do or don't occur, dynamic properties can be used for automatic monitoring. For example, a driver may be monitored, and if a number of imperfections in the driving occur over a relative short time period, this can be detected automatically. This application is similar to the one described in Sect. 13.2.1, but in this case the trace is not given at forehand, but is generated on the fly, at runtime of the real world processes. Using an automated software environment, as a form of analysis of these monitoring data it can be checked on the fly whether at some point in time some complex dynamic properties hold for these ongoing processes. A similar analysis can also be made off line. In this way the dynamic properties and the environment can be a basis for an application that addresses monitoring and analysis of complex dynamic patterns in real world contexts.

13.4 Dynamic Properties Versus Model Dynamics: Verification and Personalization

In Sect. 13.3 it was discussed how dynamic properties can be applied in relation to real world processes. Another application of dynamic properties is by relating them to a more in depth analysis of a computational model; this is addressed in the current section. In Sect. 13.2 it was already discussed how dynamic properties can be identified at forehand, before or while a model is being designed, or after a model has been designed. The current section discusses more in detail how dynamic properties can be applied in the latter situation, when a model is already given.

First, in Sect. 13.4.1 it is discussed how extensive testing of the model can be done using dynamic properties; in Sect. 13.4.2 it is discussed how characteristics of a model, represented by values of its parameters, can be personalized for a given real world situation.

13.4.1 Testing, Focusing and Analysis of a Model by Verifying It Against Dynamic Properties

Using a simulation model a large variety of traces can be generated, for example, to explore different initial values and input from the environment by systematically or randomly varying them. If these generated traces are fed into the software environment it will automatically be determined which of these traces have some specified dynamic properties. As all of this can be automated, this can support extensive forms of testing, focusing and analysis of a model, and, for example, be used to find out how the emergence of some dynamic property depends on the initial values, and external input from the environment over time, which can be used to obtain some form of predictability of emergent properties depending on such settings. Doing this, (partial) knowledge on the entailment relation between model and dynamic properties is obtained.

13.4.2 Personalizing Characteristics of a Model Based on Dynamic Properties

While generating different simulation traces, not only variations can be made for initial values and input from the environment, but also variations for the characteristics of the represented by the model's parameters (for example, connection weights, speed factors, steepness and threshold values). In this way different types of persons can be represented by the model: the model is tuned to specific situations. Partly this can be done by hand, for example by assuming that certain types of persons have strong or weak connections between some states, or higher or lower thresholds for certain states. However, it can be difficult to find appropriate values more precisely, especially when a model has many parameters. Addressing this problem can also be automated using dynamic properties that characterize a person that is to be modeled. The idea is then to generate large numbers of simulation traces for different settings of the parameter values, thereby exploring as much as possible the whole space of parameter values. This larger set of traces can be fed into the checking environment, together with the dynamic properties that are considered characteristic. The software environment will show as output which parameter settings make that the dynamic properties are satisfied and which do not make them satisfied.

13.4.3 Validation of a Model Based on Validated Dynamic Properties

The approaches discussed in Sects. 13.4.1 and 13.4.2 also provide a possibility to validate a model indirectly via dynamic properties. Then it is assumed that the dynamic properties themselves have been validated, for example as discussed in Sects. 13.3.1 or 13.3.2. Due to this assumption, the dynamic properties can be considered a valid representation of the real world's properties. Then as a second step the model can be verified and tuned so that the dynamic properties are satisfied by the model. This provides another way to validate a model by a two-step process:

- verification of the model with respect to dynamic properties
- validation of the dynamic properties.

13.5 Conceptual Representations of Dynamic Properties

Dynamic properties can be expressed in different manners and types of representation, from conceptual representations to more formal representations. In the current section conceptual representations are discussed. In Sect. 13.6 (formal) numerical-logical representations are discussed.

Conceptual representations make use of natural language expressions. Such natural language expressions could be of any form, or of a more structured form, according to some standard structure. As natural language in general can be rather ambiguous, it is better to use a structured natural language format that allows excluding ambiguity as much as possible. Such a structured natural language format for conceptual representations of dynamic properties will be introduced in this section.

The expressions for dynamic properties describing dynamic patterns have in common that they refer to states (or variables) and time points. More specifically, within these expressions references (either explicitly or implicitly) occur to:

- states
- time points
- relations between states and time points
- state relations or state properties
- temporal order relations between time points such as 'before' and 'after', 'later'
- logical relations such as 'when ...', 'and', 'or', 'not', 'eventually', 'always', 'never', 'during', 'for some time point', or 'for all time points'

For example, consider a dynamic property

When state *X* occurs, at some later point in time state *Y* occurs

This dynamic property concerns the following elements (either implicitly or explicitly expressed):

•	states	X and Y
•	time points	
•	relations between states and time points	
•	temporal relation between two time points	later
•	state relations or state properties	occurs
•	implication	when ..., ..
•	existential time quantifier	some .. point in time
•	universal time quantifier	

Sometimes some elements are left implicit in natural language expressions, such as time points, relations between states and time points, or quantifiers. In the above example, indeed these four elements do not occur explicitly in the expression; they are implicitly implied by the following expressions:

•	time points	when ..., ... at ... point in time ...
•	state relations or state properties	X occurs, Y occurs
•	relations between states and time points	when state X .. point in time state Y ...
•	universal time quantifier	when ..

Making the time points explicit by indicating them by some names, such as t , t' or t_1 , t_2 , .. provides an expression that is more clear and less ambiguous. Similarly the state properties ‘state X occurs’, and ‘state Y occurs’ can be made less ambiguous by expressing them more explicitly as ‘state X has level ≥ 0.8 ’, and ‘state Y has level ≥ 0.7 ’. These two reformulations provide a much more explicit expression for the dynamic property:

At any point in time t , when at t state X has level ≥ 0.8 ,
at some later point in time t' state Y has level ≥ 0.7

Or, using the notation $t' > t$ for ‘ t' later than t ’:

At any point in time t , when at t state X has level ≥ 0.8 ,
at some point in time $t' > t$ state Y has level ≥ 0.7

In this reformulation also the relation between states and time points has become more explicit. Finally, also the time quantifiers can be made more explicit and expressed in a standardized manner, thus making also them less ambiguous. For example, the expression ‘at any point in time t ’ can be expressed in a standardized manner by ‘for all time points t ’, and ‘at some t ’ can be expressed by ‘there exists a t ’:

Table 13.1 Elements used in conceptual representations

Element	Conceptual representation
State	$X, Y, \dots$
Time points	$t, t', t_1, t_2, \dots$
Relations between states and time points	at time t state $X \dots$
Temporal relations between time points	$t' > t, t' < t, t' \geq t, t' \leq t + d$
State relations or state properties	$\dots \geq \dots; \dots = \dots; \dots < \dots$
Implication	when $\dots, \dots$; implies $\dots, \dots$; $\dots$ implies $\dots$
Conjunction	$\dots$ and $\dots$
Disjunction	$\dots$ or $\dots$
Negation	not $\dots$
Universal time quantifier	for all time points $t \dots$
Existential time quantifier	there exists a time point $t \dots$
Universal value quantifier	for all values $V \dots$
Existential value quantifier	there exists a value $V \dots$
Universal state quantifier	for all states $X \dots$
Existential state quantifier	there exists a state $X \dots$
Universal trace quantifier	for all traces $tr \dots$
Existential trace quantifier	there exists a trace $tr \dots$

For all time points t , when at t state X has level ≥ 0.8 ,
there exists a point in time $t' > t$ such that at t' state Y has level ≥ 0.7

More in general, the standard elements used for conceptual representations of dynamic properties are as shown in Table 13.1.

A *state relation* relates states and their values to each other and to other values. Examples of state relations are:

- The level of state X at time t is at least the level of state Y at time t
- The level of state X at time t' is at most 0.05 higher than the level of state Y at time t

A state relation is *instantaneous*, when it refers to only one time point. The first state relation listed above is instantaneous, while the second one is not. A state relation involving only one state X is called a *state property* for state X ; examples of state properties are (here the first one is an instantaneous state property and the second one is not instantaneous):

- The level of state X at time t is at least 0.6
- The level of state X at time t' is at least 0.1 higher than the level of state X at time t

The above listed elements can be used to express a dynamic property in a more standard and less ambiguous way. This textual format can be structured in the following form:

For all
 time points t ,
 when
 at t state X has level ≥ 0.8 ,
 there exists
 a point in time $t' > t$ such that
 at t' state Y has level ≥ 0.7

Instead of the (prefix) expression ‘when ..., ...’, also the synonymous expression ‘... implies ...’ (infix form), or ‘implies ..., ...’ (prefix form) can be used:

For all
 time points t ,
 at t state X has level ≥ 0.8
 implies
 there exists
 a point in time $t' > t$ such that
 at t' state Y has level ≥ 0.7

Or

For all
 time points t ,
 implies
 at t state X has level ≥ 0.8 ,
 there exists
 a point in time $t' > t$ such that
 at t' state Y has level ≥ 0.7

This format makes more clear what the different parts of the expression are, and in particular in which part of the property expression a time variable occurs (the scope of the variable). It can also be depicted as a box-in-box structure as shown in Fig. 13.1.

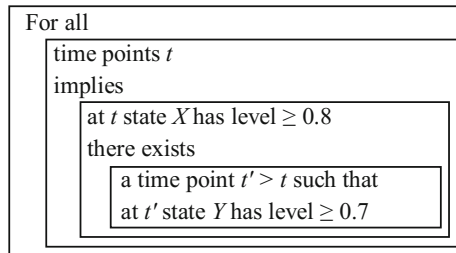


Fig. 13.1 Graphical conceptual representation of an example dynamic property as a box-in-box structure

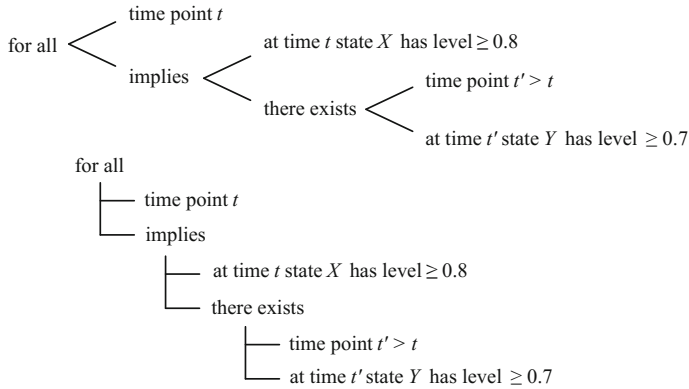


Fig. 13.2 Graphical conceptual representations of an example dynamic property as two variants of a tree-structure

These box-in-box structures are called the *nesting* of the expression. For example, the last line about state property is three levels deep in the nesting. This structure can also be depicted graphically in different variations of a tree structure, as shown in Fig. 13.2.

To illustrate this for some other examples, using these standard elements dynamic properties DP4 to DP7 from Sect. 13.2 can be expressed by a conceptual representation as follows:

DP4 For all
 time point t ,
 when
 at t execution of action a has level ≥ 0.8 ,
 there exists
 a time point $t' < t$ such that
 at t' feeling b has level ≥ 0.8

DP5 For all
 time point t ,
 when
 at t feeling b has level ≥ 0.8 ,
 there exists
 a time point $t' < t$ such that
 at t' the sensory representation of s has level ≥ 0.8

DP6 For all
 time point t ,
 at t the feeling for b has level ≤ 1

DP7 When
 for all
 time points t
 at t stimulus s in the world has level 1,
 for all
 time points t_1, t_2 with $t_2 \geq t_1$
 the level for feeling b at t_1 is at most the level for feeling b at t_2

The properties as discussed above are of different types. In Sect. 13.7 such types will be distinguished and discussed in more detail.

By expressing dynamic properties using the different elements in a standard form, conceptual representations of them are obtained, that are less ambiguous compared to arbitrary natural language expressions. Moreover, it becomes easier to transform these conceptual representations into numerical-logical representations, as is discussed in Sect. 13.6.

13.6 Numerical-Logical Representations of Dynamic Properties

In Sect. 13.5 it already became clear that expressing dynamic properties involves referring to state relations and time points. Conceptual representations have been introduced taking these into account. In the current section more formal representations are introduced: numerical-logical. Here the term ‘numerical’ concerns the way in which state relations and time points are represented based on (real) numbers. The term ‘logical’ refers to the logical structure that can be identified in a dynamic property. A numerical-logical representation provides a formalisation of both aspects, and makes it possible to use computational methods to analyse or verify dynamic properties.

13.6.1 Numerical Representations of State Relations

A state relation relates states and their values to each other and to other values. In Sect. 13.5 conceptual representations of state relations were expressed using phrases such as ‘at time t state X has level ...’. In the numerical representation of temporal-causal network models, already a numerical notation for the level of state X at time t was used, namely $X(t)$. This will also be adopted for the state relations within numerical-logical representations of dynamic properties. Using this, examples of numerical representations for state relations are:

$$\begin{aligned} X(t_2) &\leq Y(t_1) + 0.05 \\ X(t) - Y(t) &\geq 0.5 \end{aligned}$$

Similarly, examples of numerical representations for state properties are

$$\begin{aligned} X(t) &\leq 0.3 \\ X(t_2) &\leq X(t_1) + 0.02 \end{aligned}$$

State relations can involve more complex arithmetical expressions and multiple states and time points; for example, a numerical representation of the dynamic property that the values of states X and Y are always close to each other can be expressed as

$$\text{For all time points } t \text{ it holds } \text{ABS}(X(t) - Y(t)) \leq 0.05$$

A numerical representation of the dynamic property that the values of states X and Y are close to each other for multiple time points t_1 and t_2 can be expressed using, for example,

$$\sqrt{((X(t_1) - Y(t_1))^2 + (X(t_2) - Y(t_2))^2) / 2} \leq 0.05$$

In more detail, time expressions, numerical term expressions and state relations are defined as follows

Sorts

The different types of elements can be made more explicit by assuming a sorted predicate logic, for example with a sort **TIME** for time points. Then the variables indicated by $t, t', t_1, t_2, \dots$ are of sort **TIME**. Similarly sorts such as **REAL**, **STATE** and **TRACE** can be used, each with the variable notation (possibly with subscripts or primes) shown and with equality relation $=$, and **TIME** and **REAL** also with an ordering relation $<$:

variable	sort	contains
t	TIME	time points
V	REAL	real numbers
X, Y	STATE	states
tr	TRACE	traces

To be able to check a dynamic property against a trace, each sort needs to contain only a finite number of instances (constants for names). For the sort TIME, this finite number of time points can be derived from the trace that is used, which itself only contains a finite number of time points. For the sort REAL this has to be specified, for example, by using the real numbers within the interval $[0, 1]$ in three decimals. For the sort STATE the instances are the states (names) actually used in the model that is considered, and for the sort TRACE the finite number of traces (names) considered.

Time representations

Time representations are numerical expressions of sort TIME and can be of the following types:

1. t is a time expression for each constant or variable t for a time point; time points are real numbers
2. $t + D$ and $t - D$ are time expressions where t is a constant or variable for a time point and D a constant or variable for a real number (sort VALUE)

In this definition, time expressions are kept simple. If more complex time expressions are required, item 2. easily can be extended to more complex numerical expressions.

Numerical term representations

Numerical term representations are expressions of sort VALUE built inductively as follows:

1. V is a numerical term representation, for each constant or variable V for a real number
2. te is a numerical term representation for each time expression te
3. $X(t)$ is a numerical term representation for any state name X , and time expression t
4. $V(tr, X, t)$ is a numerical term representation for any constant or variable for state X , trace tr , and any time expression t
5. $f(nte_1, nte_2)$ and $g(nte_1)$ are numerical term representation for any numerical term representations nte_1 and nte_2 and any binary arithmetical functions $f : \text{VALUE} \times \text{VALUE} \rightarrow \text{VALUE}$ and unitary arithmetical functions $g : \text{VALUE} \rightarrow \text{VALUE}$.

Note that here in 3. and 4. alternative notations are used for the value of a state X at time t ; here for each state X a function $X : \text{TIME} \rightarrow \text{VALUE}$ is assumed, and as

an alternative one more generic function $V: \text{TRACE} \times \text{STATE} \times \text{TIME} \rightarrow \text{VALUE}$ is assumed. They indicate the same values, but the numerical term expressions $V(tr, X, t)$ in 4. enables to refer explicitly to state X and/or trace tr and to use quantifiers over states and traces (this is sometimes called reification). The expression $V(tr, X, t)$ has to be used when variables and quantifiers over multiple traces and/or states are involved; see Sect. 13.7.3 for types of dynamic properties in which this happens. In all other cases also just the representations described in 3. can be used for the value of a state X at time t . In 5. the standard arithmetical functions can be used, such as $+$, $-$, $*$, $/$, 2 , $\sqrt{}$, By 5. in combination with 2. and 3. it is possible to express, for example, the speed of change $(X(t_2) - X(t_1))/(t_2 - t_1)$ of a state X from time t_1 to t_2 .

Numerical state relation representations

State relation representations are expressions of the form $R(nte_1, nte_2)$ where nte_1 and nte_2 are numerical term representations and R is a binary relation, taken from $=, <, >, \leq, \geq$.

Examples of state relations are $a_1 < a_2$, $a_1 = a_2$, $a_1 \leq a_2$, $a_1 > a_2$, $a_1 \geq a_2$ where a_1, a_2 denote state terms such as $V(tr_2, X_1, t)$, or $X_2(t)$, variables or constants such as ε or an arithmetical expression using such elements, such as $\text{ABS}(X(t_2) - \underline{e}_X)$, also denoted as $|X(t_2) - \underline{e}_X|$.

13.6.2 Using Numerical Representations Within a Dynamic Property Expression

Using the numerical representations $X(t)$ and $Y(t')$ for the relation between states and time points, the dynamic property considered conceptually in Sect. 13.5 can also be written as follows:

For all
 time points t ,
 when
 $X(t) \geq 0.8$,
 there exists
 a point in time $t' > t$ such that
 $Y(t') \geq 0.7$

Instead of the prefix expression ‘when ..., ...’, also the synonymous infix expression ‘... implies ...’, or prefix expression ‘implies ..., ...’ can be used:

For all
 time points t ,
 $X(t) \geq 0.8$
 implies
 there exists
 a point in time $t' > t$ such that
 $Y(t') \geq 0.7$

Or

For all
 time points t ,
 implies
 $X(t) \geq 0.8$,
 there exists
 a point in time $t' > t$ such that
 $Y(t') \geq 0.7$

Moreover, the properties DP4 to DP7 from Sects. 13.2 and 13.5 can also be expressed as:

DP4 For all
 time point t ,
 $es_a(t) \geq 0.8$,
 implies
 there exists
 a time point $t' < t$ such that
 $fs_b(t') \geq 0.8$

DP5 For all
 time point t ,
 $fs_b(t) \geq 0.8$,
 implies
 there exists
 a time point $t' < t$ such that
 $srs_s(t') \geq 0.8$

- DP6

For all
time point t ,
 $fs_b(t) \leq 1$
- DP7

For all
time points t
 $ws_s(t) = 1$,
implies
for all
time points t_1, t_2 with $t_2 \geq t_1$
 $fs_b(t_1) \leq fs_b(t_2)$

13.6.3 Numerical-Logical Representation of a Dynamic Property Expression

Usually a dynamic property involves a number of state relations, but these are embedded in a structure of temporal and logical relations between them. For example, consider a conditional achievement property:

For all
time points t ,
when
 $X(t) \geq 0.8$,
there exists
a point in time $t' > t$ such that
 $Y(t') \geq 0.7$

At the basic level this dynamic property involves the two state relations $X(t) \geq 0.8$ and $Y(t') \geq 0.7$. But in addition it involves the following logical relations between these state relations; these can be represented by a formal logical representation as shown in the third column:

•	universal time quantifier relation	for all time points t	$\forall t$
•	implication relation	when ..., ...	$\rightarrow$
•	existential time quantifier relation	there exists a time point t'	$\exists t'$

Table 13.2 Conceptual and formal representations for logical relations

Logical relation	Conceptual representation	Formal logical representation
Implication	when ..., ...; implies ..., ...; ... implies ...	$\dots \rightarrow \dots$
Conjunction	... and ...	$\dots \wedge \dots$
Disjunction	... or ...	$\dots \vee \dots$
Negation	not ...	$\neg \dots$
Universal time quantifier	for all time points t ...	$\forall t \dots$
Existential time quantifier	there exists a time point t ...	$\exists t \dots$
Universal value quantifier	for all values V ...	$\forall V \dots$
Existential value quantifier	there exists a value V ...	$\exists V \dots$
Universal state quantifier	for all states X ...	$\forall X \dots$
Existential state quantifier	there exists a state X ...	$\exists X \dots$
Universal trace quantifier	for all traces tr ...	$\forall tr \dots$
Existential trace quantifier	there exists a trace tr ...	$\exists tr \dots$

So, a logical implication relation ‘when ..., .., or ...’, ‘implies ..., ...’ is expressed formally as $\rightarrow$, and universal and existential quantifiers by $\forall$ and $\exists$, respectively. Then, the (formal) numerical-logical representation for the above dynamic property is

$$\forall t_1 [X(t_1) \geq 0.8 \rightarrow \exists t_2 [t_2 \geq t_1 \wedge Y(t_2) \geq 0.7]]$$

This is a formal expression in temporal predicate logic, using ordering relations such as $\geq$ for time points and values, and function symbols $X(t)$ from time points to values for each state X . Notice the use of [and], or (and), to indicate what belongs together, and in particular what falls within the scope of a quantifier. In Table 13.2 an overview is shown of different logical relations and their conceptual and formal representations.

Numerical-logical dynamic property representations

Numerical-logical dynamic property representations are defined inductively on top of the numerical state relation representations as follows

1. Each numerical state relation representation is a numerical-logical dynamic property representation
2. If DP and DP' are numerical-logical dynamic property representations then

$$DP \wedge DP', DP \vee DP', DP \rightarrow DP', \neg DP$$

are numerical-logical dynamic property representations

3. If DP is a numerical-logical dynamic property representation, then

$$\forall t \ DP, \exists t \ DP, \forall V \ DP, \exists V \ DP$$

are numerical-logical dynamic property representations, where t is a variable over time points, and V over values.

4. If DP is a numerical-logical dynamic property representation in which no numerical expression $Y(t)$ occurs for any Y and t , but only numerical expressions of the form $V(tr, Y, t)$, then

$$\forall X \ DP, \exists X \ DP, \forall tr \ DP, \exists tr \ DP$$

are numerical-logical dynamic property representations, where X is a variable over states, and tr over traces. Note that, when one of $\forall X$, $\exists X$, $\forall tr$, or $\exists tr$ is to be applied, and in DP there still are occurrences of $Y(t)$ for Y and t , these occurrences can be replaced by $V(tr, X, t)$ after which a quantifier X or tr can be applied.

Usually some abbreviations are used:

$\exists t_2 \geq t_1 \ A$	abbreviates	$\exists t_2 [t_2 \geq t_1 \wedge A]$
$\forall t_2 \geq t_1 \ A$	abbreviates	$\forall t_2 [t_2 \geq t_1 \rightarrow A]$

for any property A , and similarly for other ordering relations and quantifiers. For multiple occurrences of a universal or an existential quantifier usually the following abbreviation is used:

$\forall t_1, t_2 \ A$	abbreviates	$\forall t_1 \forall t_2 \ A$
$\exists t_1, t_2 \ A$	abbreviates	$\exists t_1 \exists t_2 \ A$
$\forall t_1, X \ A$	abbreviates	$\forall t_1 \forall X \ A$

Both abbreviations can be combined, for example

$\exists t_1, t_2 > t_1 \ A$	abbreviates	$\exists t_1, t_2 [t_2 > t_1 \ \& \ A]$
$\forall t_1, t_2 \geq t_1 \ A$	abbreviates	$\forall t_1, t_2 [t_2 \geq t_1 \rightarrow A]$

In this way the dynamic properties DP4 to DP7 from Sects. 13.2 and 13.5 can be expressed by numerical-logical representations as follows:

DP4 $\forall t [\text{es}_a(t) \geq 0.8 \rightarrow \exists t' < t \ \text{fs}_b(t') \geq 0.8]$

DP5 $\forall t [\text{fs}_b(t) \geq 0.8 \rightarrow \exists t' < t \ \text{srs}_s(t') \geq 0.8]$

DP6 $\forall t \text{ fs}_b(t) \leq 1$

DP7 $[\forall t \text{ ws}_s(t) = 1] \rightarrow [\forall t_1, t_2 \geq t_1 \text{ fs}_b(t_1) \leq \text{fs}_b(t_2)]$

13.7 Types of Dynamic Properties and Their Representations

In this section a number of specific types of dynamic properties are discussed. Note that not every possible dynamic property is of one of these types, and the different types of properties are not assumed to be disjoint. In Tables 13.1, 13.2 and 13.3 it is shown how they can be expressed in conceptual and formal logical representations.

13.7.1 Basic State Relation, Achievement, Grounding, Representation, Ordering and Monotonicity Properties

The first types of dynamic properties discussed relate the occurrence of states at different time points with longer time periods between them: achievement properties, grounding properties, representation properties, ordering properties, and monotonicity properties.

Basic state relation properties

A first type of dynamic properties considered concerns basic state relations. They express direct relations for levels of specific states at specific instances of time points. For example:

- At time point 20 the level of state *Y* is higher than the level of state *X* at time point 10

Basic state relation properties are expressed in terms of ground state relations; no variables or quantifiers are used. In that sense they are rather simple. But they can be quite relevant when a model is compared to a real world pattern for which empirical data on specific state levels at specific time points are given. For example, suppose empirical data are available for state instances $\underline{X}_1, \dots, \underline{X}_M$ and instances of time points $\underline{t}_1, \dots, \underline{t}_N$, by values $\underline{V}_{i,j}$ for these states $\underline{X}_i$ at these time points $\underline{t}_j$. In general such empirical data can be represented as a set of triples $(\underline{X}_i, \underline{t}_j, \underline{V}_{i,j})$ for $i = 1, 2, \dots, M$ and $j = 1, 2, \dots, N$, where $\underline{X}_i, \underline{t}_j, \underline{V}_{i,j}$ refer to ..specific instances of states, time points and values:

$$(\underline{X}_1, \underline{t}_1, \underline{V}_{1,1}), \dots, (\underline{X}_1, \underline{t}_N, \underline{V}_{1,N}),$$

$$(\underline{X}_M, \underline{t}_1, \underline{V}_{M,1}), \dots, (\underline{X}_M, \underline{t}_N, \underline{V}_{M,N})$$

Then the dynamic property

$$\begin{aligned} \mathbf{DP} \quad V(tr, \underline{X}_1, t_1) &= \underline{V}_{1,1} \wedge \dots \wedge V(tr, \underline{X}_1, t_N) = \underline{V}_{1,N} \wedge \\ V(tr, \underline{X}_M, t_1) &= \underline{V}_{M,1} \wedge \dots \wedge V(tr, \underline{X}_M, t_N) = \underline{V}_{M,N} \end{aligned}$$

is a basic state relation property expressing that in trace tr (for example generated by simulation) for each i and j the value of state $\underline{X}_i$ at time point t_j is equal to $\underline{V}_{i,j}$. This can also be expressed in a short notation

$$\wedge_{i,j} V(tr, \underline{X}_i, t_j) = \underline{V}_{i,j}$$

where $\wedge_{i,j}$ denotes the conjunction for all i and j .

Such a dynamic property DP may not hold in a strict sense for a given simulation trace tr , but it may be found out that it is *fulfilled approximately* in the sense that the values of states $\underline{X}_i$ at these time points t_j in the simulation trace tr may not be exactly the same as the empirical values, but are also not that different from them. This can be covered by expressing (slightly) less strict variants of dynamic property DP. For example, the following less strict variant of this basic state relation property DP may hold, in which case it asserts that DP holds approximately, with accuracy D ; it expresses that the absolute values of the differences are at most a given small positive number D :

$$\begin{aligned} \mathbf{DP}(D) \quad |V(tr, \underline{X}_1, t_1) - \underline{V}_{1,1}| &\leq D \wedge \dots \wedge |V(tr, \underline{X}_1, t_N) - \underline{V}_{1,N}| \leq D \wedge \\ |V(tr, \underline{X}_M, t_1) - \underline{V}_{M,1}| &\leq D \wedge \dots \wedge |V(tr, \underline{X}_M, t_N) - \underline{V}_{M,N}| \leq D \end{aligned}$$

or in short notation

$$\wedge_{i,j} |V(tr, \underline{X}_i, t_j) - \underline{V}_{i,j}| \leq D$$

In Chap. 14 it will be pointed out how this can be exploited by heuristic search methods to find appropriate parameter values to approximate the empirical data.

Achievement properties

Next, properties are considered that indicate that eventually a process reaches a kind of final or intermediate state in which something has been *achieved*. In contrast to basic state relation properties as described above, here the time point at which it is achieved is not prescribed. An achievement property for the simple example model from Chap. 2 is:

DP1 When stimulus s occurs, action a will occur

Other properties are

- After a good feeling about its predicted effect the action will occur
- After the stimulus and a good feeling about its predicted effect the action will occur

In a conditional form an achievement property can be expressed as ‘always if state X occurs, then at some later time state Y occurs’. For the case that some intermediate state within a longer lasting process is reached, this type of property can also be considered a milestone property. Such properties can be used, for example, to analyse the overall process as a number of sub-processes that, maybe depending on other milestones already achieved, each result in the achievement of some milestone. Here Y is a milestone and X can be a preceding milestone.

A time-constrained achievement property has a built-in limitation on the timing of the action:

- When stimulus s occurs, action a will occur within 10 s

When intermediate states within a longer lasting process are reached, that can also be expressed as achievement properties. As discussed in Sect. 13.2 for the example model the sensory representation of s and feeling b could be considered milestones, first from the occurrence of s to sensory representation of s (DP2), then from sensory representation of s to feeling b (DP3), and finally, from feeling b to performing action a (DP4), expressed by the following three milestone properties:

- DP2** Always, when stimulus s occurs in the world, at some point in time after this a sensory representation of s will occur
- DP3** Always, when a sensory representation of s occurs, at some point in time after this feeling b will occur
- DP4** Always, when feeling b occurs, at some point in time after this execution of action a will occur

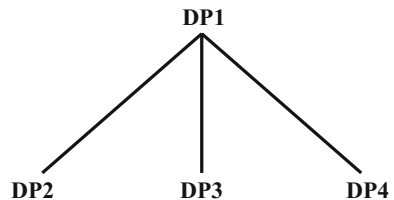
These three dynamic properties DP2, DP3, DP4 can be considered a *refinement* of the overall achievement property DP1; e.g., (Herlea et al. 1999, 2005; Ferber et al. 2001). A conceptual representation of such a refinement can be graphically depicted as shown in Fig. 13.3.

A refinement can also be expressed by the following logical relation:

$$\text{DP2} \ \& \ \text{DP3} \ \& \ \text{DP4} \Rightarrow \text{DP1}$$

This refinement relation and the associated logical relation show that the requirement for the model as a whole can be satisfied when the model is designed in such a manner that it satisfies refined requirements DP2, DP3 and DP4. As

Fig. 13.3 Refinement relation for a dynamic property



discussed in Sect. 13.2 in this way such a refinement can provide a focus for the design process by designing the model so that the refined requirements DP2, DP3 and DP4 are fulfilled.

Grounding properties

Grounding properties point at a necessary condition for some achieved state at some point in time t , which should have occurred before that time point. Some grounding properties for the simple example model from Chap. 2 are:

- *No action without a good feeling about it:*
if an action occurs, then earlier there was a good feeling about its predicted effect
- *No feeling before a partial response:*
a feeling does not occur before a partial preparation occurs
- *No full preparation without a good feeling:*
a full preparation does not occur before a feeling occurs about its predicted effect

Representation properties

Another type of properties considered is formed by *representation properties*. These are properties that express how an internal (mental) state relates to external states (or sensor and/or execution states) in past and/or future times; e.g., (Kim 1996, pp. 184–185, 191–194, 200–202). In particular, the notion of representation based on *relational specification* as discussed in (Kim 1996, pp. 200–202) is adopted here. This considers the representational content of a mental state as being based on the specification of a representation relation between this mental state and certain states of affairs in the world distant in space and/or time. Such states of affairs can be described and formalised as a state property or a dynamic property, which is used as a component of a dynamic property describing the (temporal) relationship of this state of affairs to the mental state. For example, a state of affairs can be an observed running horse in a field, which disappears in the fog. After this, the belief that there is a horse in the field is a mental state that relates to the occurrence of this pattern of a running horse in the recent past. Two specific types can be distinguished, looking backward in time, or looking forward in time:

Backward representation properties

These are properties that relate an internal state backward in time to external world states or sensor or execution states in the past; for example, some mental state (e.g., a representation state or a preparation state) has a high activation level when a stimulus occurred before or has been sensed before.

Forward representation relations

These are properties that relate an internal state forward in time to external world states or sensor or execution states in the future; for example, if a preparation state has a high activation level, then later on the effect of the corresponding action will be found in the world.

Examples of representation properties for the simple example model from Chap. 2 are:

- *Action a always has effect e*
If a preparation state for action a occurs, then later the effect e will occur in the world.
- *Always after the occurrence of a stimulus a sensory representation occurs*
If a sensory representation of a stimulus s occurs, then earlier this stimulus s occurred in the world, and if a stimulus s in the world occurs, then later a sensory representation of this stimulus s will occur.

Ordering properties

Ordering properties express an ordering relation between the occurrence of one state property and another one: state property p occurs before state property q .

An example for the simple example model from Chap. 2 is:

- The order of occurrence of states is: stimulus in world state —sensor state—sensory representation state—partial preparation state—feeling state—high level preparation state—execution state.

Monotonicity properties

These are properties expressing, for example, that under certain conditions (e.g., after some time point) the value of a variable is *monotonically increasing* or *decreasing* over time. They indicate that the activation level of a state X increases over time, or the activation level of state X decreases over time. Such properties can be expressed as ‘if $t < t'$ then $X(t) < X(t')$ ’, perhaps preceded by some conditions. They also have been addressed in Chap. 12. Examples for the simple example model from Chap. 2 are:

- *Constant stimulus level 1 makes that all states will have increasing levels:*
If from some time point on a constant stimulus with level 1 occurs, then from some later time point on all states will have monotonically increasing levels
- *Constant stimulus level 0 makes that all states will have decreasing levels:*
If from some time point on a constant stimulus with level 0 occurs, from some later time point on all states will have monotonically decreasing levels

In Table 13.3 temporal predicate logic formalisations are shown for a number of types of dynamic properties: achievement, grounding, ordering, and monotonicity properties.

13.7.2 Maintenance, Peak, Speed, Equilibrium and Limit Cycle Properties

Other types of properties addressed in this section concern ongoing trends over longer time intervals: maintenance properties, peak properties, speed properties, stationary point properties, and limit cycle properties.

Table 13.3 Temporal predicate logic formalisations for dynamic properties: achievement, grounding, ordering, and monotonicity properties

Example conceptual representations	Example numerical-logical representations
Basic state relation property A property expressing state relations for specific state instances and specific instances of time points	
In trace tr for each i and j the value of state $\underline{X}_i$ at time point t_j is equal to $\underline{V}_{i,j}$	$\begin{aligned} V(tr, \underline{X}_1, t_1) = \underline{V}_{1,1} \wedge \dots \wedge V(tr, \underline{X}_1, t_N) = \underline{V}_{1,N} \\ \wedge \dots \wedge \\ V(tr, \underline{X}_M, t_1) = \underline{V}_{M,1} \wedge \dots \wedge V(tr, \underline{X}_M, t_N) = \underline{V}_{M,N} \end{aligned}$
In trace tr for each i and j the value of state $\underline{X}_i$ at time point t_j differs from $\underline{V}_{i,j}$ by at most a given (small) positive number D	$\begin{aligned} V(tr, \underline{X}_1, t_1) - \underline{V}_{1,1} \leq D \wedge \dots \\ \wedge V(tr, \underline{X}_1, t_N) - \underline{V}_{1,N} \leq D \wedge \\ \dots \wedge \\ V(tr, \underline{X}_M, t_1) - \underline{V}_{M,1} \leq D \wedge \dots \\ \wedge V(tr, \underline{X}_M, t_N) - \underline{V}_{M,N} \leq D \end{aligned}$
Achievement property Given no or some conditions, eventually a certain state is reached in which some state property holds	
There exists a time point t such that at t state $Y(t)$ has level $\geq a$	$\exists t X(t) \geq a$
For all time points t_1 when at t_1 state X has level $\geq a_1$, there exists a time point $t_2 \geq t_1$ such that at t_2 state Y has level $\geq a_2$	$\forall t_1 [X(t_1) \geq a_1 \rightarrow \exists t_2 \geq t_1 Y(t_2) \geq a_2]$
Grounding property When a state property q occurs, state property p has occurred before this	
For all time points t_2 , when at t_2 state Y has level $\geq a_2$, there exists a time point $t_1 \leq t_2$ such that at t_1 state X has level $\geq a_1$	$\forall t_2 [Y(t_2) \geq a_2 \rightarrow \exists t_1 \leq t_2 X(t_1) \geq a_1]$
Representation property <i>Backward</i> : relating an internal state backward in time to external world states or sensor or execution states in the past. <i>Forward</i> : relating an internal state forward in time to external world states or sensor or execution states in the future	
For all time points t_1 , when at t_1 stimulus state S has level $\geq a_1$, then there exists a time point $t_2 \geq t_1$ such that at t_2 internal state R has level $\geq a_2$, and conversely	$\begin{aligned} \forall t_1 [S(t_1) \geq a_1 \rightarrow \exists t_2 \geq t_1 R(t_2) \geq a_2] \\ \forall t_2 [R(t_2) \geq a_2 \rightarrow \exists t_1 \leq t_2 S(t_1) \geq a_1] \end{aligned}$
For all time points t_1 , when at t_1 preparation state P has level $\geq a_1$, there exists a time point $t_2 \geq t_1$ such that at t_2 execution state E for the action has level $\geq a_2$, and conversely	$\begin{aligned} \forall t_1 [P(t_1) \geq a_1 \rightarrow \exists t_2 \geq t_1 E(t_2) \geq a_2] \\ \forall t_2 [E(t_2) \geq a_2 \rightarrow \exists t_1 \leq t_2 P(t_1) \geq a_1] \end{aligned}$
Ordering property State X occurs before state Y	
There exist time points t_1 and t_2 with $t_1 < t_2$ such that at t_1 state X has level $\geq a_1$ and at t_2 state Y has level $\geq a_2$	$\exists t_1, t_2 > t_1 [X(t_1) \geq a_1 \ \& \ Y(t_2) \geq a_2]$
Monotonicity property The activation level of state X increases over time; the activation level of state X decreases over time	
For all time points t_1 and $t_2 > t_1$, the level of state X at t_2 is higher than the level of state X at t_1	$\forall t_1, t_2 > t_1 X(t_2) > X(t_1)$
For all time points t_1 and $t_2 > t_1$, the level of state X at t_2 is lower than the level of state X at t_1	$\forall t_1, t_2 > t_1 X(t_2) < X(t_1)$

Maintenance properties

These are properties expressing that during a process (or part thereof) a state is *always satisfying a certain state criterion*. For example, specifying that a form of homeostasis is maintained: a state in which variables stay within certain fixed bounds. A specific type of maintenance properties express that a certain state level of a model always remains within a certain interval. These dynamic properties can be expressed as ‘for all time points t state property p holds’, perhaps preceded by some condition. Some maintenance properties for the simple example model from Chap. 2 are:

- *Stimulus levels between 0 and 1 will imply sensor and sensory representation state levels between 0 and 1:*

When the stimulus always stays within the interval from 0 to 1, then the activation levels of the sensor state and the sensory representation state of the stimulus also always stay within the interval from 0 to 1.

- *Sensory representation state levels between 0 and 1 will imply preparation state levels between 0 and 1:*

When the sensory representation of the stimulus always stays within the interval from 0 to 1 (and the sum of the weights of the connections to the preparation state is at most 1), then the activation levels of the preparation state and the execution state of the action also always stay within the interval from 0 to 1.

In Table 13.4 some other types of maintenance are addressed. An *upward bound* (by a number a_2) property can be expressed by a universal quantifier: for all time points t it holds $X(t) \leq a_2$. In a formal temporal predicate logic format this is expressed as $\forall t X(t) \leq a_2$. Similarly for a downward bound: $\forall t X(t) \geq a_1$. These can easily be used to express maintenance properties. In Table 13.4 it is also shown how peak properties, stationary point properties and attracting equilibrium properties can be expressed.

Peak properties

Peak properties indicate the maximal values of activation level of a state that is reached: state X reaches a maximal value of at least V , state X reaches a maximal value of at most V , the maximal value of state X is V . A special case is a *single peak property*: the activation level of state X increases to some maximum value and decreases after that time point. An example single peak property for the simple example model from Chap. 2 is:

- *The feeling level has one peak:*

The feeling level increases until some value after which it decreases

Speed properties

Speed properties indicate the speed by which a state changes over a certain time interval. For the speed of state X from time t_1 to t_2 a calculation can be used of the difference quotient $(X(t_2) - X(t_1))/(t_2 - t_1)$.

- *The speed of generating a feeling over some time interval:*

The feeling level increases from time t_1 to t_2 with speed at least v

Table 13.4 Temporal predicate logic formalisations for dynamic properties: bound, maintenance, peak, stationary point and attracting stationary point properties

Example conceptual representations	Example numerical-logical representations
Bound property State X is upward bounded by a_2 ; its maximal value is at most a_2 . State X is downward bounded by a_1 ; its minimal value is at least a_1	
For all time points t , at t state X has level $\leq a_2$ For all time points t , at t state X has level $\geq a_1$	$\forall t X(t) \leq a_2$ $\forall t X(t) \geq a_1$
Maintenance property During a process a state level always stays within certain fixed bounds	
For all time points t the level of state value X at t is at least a_1 and at most a_2 . For all time points t the levels of state X are within the interval $[0, 1]$	$\forall t [X(t) \geq a_1 \ \& \ X(t) \leq a_2]$ $\forall t [X(t) \geq 0 \ \& \ X(t) \leq 1]$
Peak property The activation level of state X increases to some maximum value and decreases after that time point	
There exists a time point t such that for all time points t_1 and t_2 with $t_1 \leq t_2 < t$ the level of state X at t_1 is lower than the level of state X at t_2 , and for all time points t_1 and t_2 with $t \leq t_1 < t_2$ the level of state X at t_1 is higher than the level of state X at t_2	$\exists t [\forall t_2 \leq t, t_1 < t_2 X(t_1) < X(t_2)]$ $\wedge \forall t_1 \geq t, t_2 > t_1 X(t_1) > X(t_2)]$
Speed property The feeling level increases from time t_1 to t_2 with speed at least v	
The difference quotient for the feeling state from time t_1 to t_2 is at least v The difference quotient for the feeling state from time t to $t + \varepsilon$ is at least v	$(X(t_2) - X(t_1))/(t_2 - t_1) \geq v$ $(X(t + \varepsilon) - X(t))/\varepsilon \geq v$
Stationary point property A state in which values of variables do not change over time An stationary point property can be expressed as a kind of conditional maintenance property	
For all time points t_1 , when at t_1 state X has level $\underline{e}_X$, for all time points $t_2 > t_1$, at t_2 state X has level $\underline{e}_X$	$\forall t_1 [X(t_1) = \underline{e}_X \rightarrow \forall t_2 > t_1 X(t_2) = \underline{e}_X]$ for all states X
Attracting stationary point property Such a property expresses that eventually an (attracting) stationary point is approximated	
There exists a point in time t_1 such that for all time points $t_2 \geq t_1$, at t_2 the difference between the level of state X and $\underline{e}_X$ is less than some small positive number ε (e.g., $\varepsilon = 0.0001$)	$\exists t_1 \forall t_2 \geq t_1 X(t_2) - \underline{e}_X < \varepsilon$ for all states X
Limit cycle property All variables follow a fixed periodically recurring pattern over time	
There is a time duration D such that for all states X and all time points t the level of state X at $t + D$ is the same as the level of state X at t	$\exists D \forall t X(t + D) = X(t)$ for all states X

Note that also the instantaneous speed at one time point t can be considered. This can be done to choose a small $\varepsilon > 0$ (for example, the step size of a simulation trace or data set) and consider the speed from t to $t + \varepsilon$.

- *The instantaneous speed of generating a feeling at some point in time:*
The feeling level increases at time t with instantaneous speed at least v

Note that when a dynamic model is given, also from the differential equation for the state that is considered, an expression for the instantaneous speed can be obtained.

Stationary point properties

A stationary point property with stationary point value $\underline{e}$ for state X can be expressed as a kind of conditional maintenance property: ‘if for some time point t state property $X(t) = \underline{e}$, then $X(t') = \underline{e}$ holds for all time points $t' > t$ ’. Note that usually such a property only holds for a collection of states, not for single states: ‘if for some time point t state properties $X_i(t) = \underline{e}_i$ for all i , then $X_i(t') = \underline{e}_i$ holds for all i and all time points $t' > t$ ’. Some stationary point state properties for the simple example model from Chap. 2 are:

- *Stimulus constant 1 implies value 1 for all states*
If a constant stimulus 1 occurs in the world, then the values 1 for all states are an equilibrium state
- *Stimulus constant 0 implies value 0 for all states*
If a constant stimulus 0 occurs in the world, then the values 0 for all states are an equilibrium state

Attracting properties

These properties express that eventually an (attracting) equilibrium is approximated. That the equilibrium is attracting can be expressed as a kind of achievement property: ‘there is a point in time t such that at t it holds $X(t)$ is very close to $\underline{e}$ ’. Note that here *very close* can be expressed as $X(t)$ having a difference with $\underline{e}$ less than some very small number such as some small positive number ε , for example, 0.001. Some attracting properties for the simple example model from Chap. 2 are:

- *Stimulus constant 1 implies for all states the levels are attracted to 1*
If from some time point on a constant stimulus with level 1 occurs, then from some later time point on all states will have a level very close to 1
- *Stimulus constant 0 implies for all states the levels are attracted to 0*
If from some time point on a constant stimulus with level 0 occurs, then from some later time point on all states will have a level very close to 0

Limit cycle properties

Sometimes a process does not end up in an equilibrium state but in a fixed periodically recurring pattern, called a limit cycle. Such a limit cycle is defined by the time duration D of this period, and is expressed by:

- *Limit cycle with period D:*
for each state X and each time point t the value of X at $t + D$ is the same as the value of X at t .

13.7.3 *State Comparison, Trace Comparison and Trace Selection Properties*

In this section dynamic properties are discussed in which multiple states or multiple traces play a role and quantifiers over traces and states can be used. In such cases the names of states X and traces tr have to be made explicit in a form that variables can be used for them. This means that numerical expressions of the type $X(t)$ are not used, but instead expressions of the form $V(tr, X, t)$. Types of properties discussed are state comparison properties, trace comparison properties and trace selection properties.

State comparison properties

State comparison properties are properties in which states are compared to a set of or all other states using quantifiers over states. One simple example is the type of dynamic property that expresses that all states satisfy a same type of property, for example monotonicity:

- All states X are monotonically increasing from all time points t to all time points $t' > t$

This can be expressed as the following conceptual representation:

- For all states X and all times points $t, t' > t$ it holds $V(tr, X, t) < V(tr, X, t')$

Another type of state comparison property is to identify the state that reaches the highest value, compared to other states:

- State X has a value at some time t such that for all t' all states Y have at most that value at t'

Another example is a dynamic property that expresses that control states are mutually exclusive; only one of them can have a higher level:

- For all control states X, Y and all time points t , if $X \neq Y$ and X has level > 0.6 at t , then Y has level < 0.4 at t

Yet another example expresses that no preparation state can occur unless at least one sensory representation has occurred:

- For all preparation states X and all time points t , if X has level > 0.6 at t , then there exists a sensory representation state Y and a time point $t' < t$ such that Y has level > 0.6 at t'

Trace comparison properties

Trace comparison properties are properties in which whole traces (trajectories, scenarios) over time are compared. These properties express, for example, that if in one trace the values of certain states are *lower* than in a second trace, then the value of some other state will also be *lower* than in that second trace. A specific example is the property that when experiences with someone in one trace are better than in a second trace, then trust in that person will be higher in that trace than in the second trace. Another example is that when in one trace more exercising takes place than in another trace, then in this former trace the skill level will be higher. Such trace comparison properties are relevant in particular for cases of adaptive or learning processes, for example, involving Hebbian learning, or dynamic network connections describing evolving social interaction.

Within such properties not only references to time points occur but also to traces. A simple trace comparison property expresses that some property holds for all traces in a given set of traces:

- For all traces tr state X increases monotonically from all time points t to all time points $t' > t$

This can be expressed by the following conceptual representation:

- For all traces tr and all times points $t, t' > t$ it holds $V(tr, X, t) < V(tr, X, t')$

This can also be expressed not for one specific state X but for all states X :

- For all traces tr all states X increase monotonically from all time points t to all time points $t' > t$

This can be expressed by the following conceptual representation:

- For all traces tr , all states X and all times points $t, t' > t$ it holds $V(tr, X, t) < V(tr, X, t')$

Another example trace comparison property, for the simple example model from Chap. 2 is:

- *Higher stimulus levels imply higher mental state levels:*

If in one trace the stimulus level in the world is always at least as high as in another trace, then for each state the level will always be at least as high in the former trace as in the latter trace

Such properties can be expressed, for example by ‘for any two traces tr_1 and tr_2 if at all time points t in trace tr_2 it holds $V(tr_1, X, t) < V(tr_2, X, t)$ (e.g., more exercising X occurs than in trace tr_2), then at all time points t it holds $V(tr_1, Y, t) < V(tr_2, Y, t)$ (e.g., in tr_2 a higher skill level Y will occur than in trace tr_1)’. Also the inverse variation is possible, for example,

- *More support leads to less stress*

When always more support is given to a person then this person will always have a lower level of stress

Table 13.5 Temporal predicate logic formalisations for dynamic properties: state and trace comparison properties

Example conceptual representations	Example numerical-logical representations
State comparison property These properties compare different states and for this use quantifiers over states, for example, the state that reaches the highest level of all states	
For all states X and all times points $t, t' > t$ it holds $V(tr, X, t) < V(tr, X, t')$	$\forall X, t, t' > t \ V(tr, X, t) < V(tr, X, t')$
There exists a time point t such that for all states Y and all time points t' the level of Y at t' is at most the level of X at t	$\exists t \ \forall Y, t' \ V(tr, Y, t') \leq V(tr, X, t)$
For all control states X, Y and all time points t , if $X \neq Y$ and X has level > 0.6 at t , then Y has level < 0.4 at t	$\forall X:CS, Y:CS, t \ [X \neq Y \wedge V(tr, X, t) > 0.6] \rightarrow V(tr, Y, t) < 0.4]$
For all preparation states X and all time points t , if X has level > 0.6 at t , then there exists a sensory representation state Y and a time point $t' < t$ such that Y has level > 0.6 at t'	$\forall X:PS, t \ [V(tr, X, t) > 0.6 \rightarrow \exists Y:SRS, t' < t \ V(tr, Y, t') > 0.6]$
Trace comparison property These properties express, for example, that if in one trace the values of certain states are <i>lower</i> than in a second trace, then the value of some other state will also be <i>lower</i> in that trace. Such trace comparison properties are relevant, for example, for cases of adaptive processes, such as for Hebbian learning or dynamic connections in a network describing evolving social interaction	
For all traces tr_1 and tr_2 , when for all time points t in trace tr_2 the level of state X at t is at least the level of state X at t in trace tr_1 , then for all time points t in tr_2 the level of Y at t is at least the level of state Y at t in trace tr_1	$\forall tr_1, tr_2$ $[\forall t \ V(tr_1, X, t) \leq V(tr_2, X, t) \rightarrow \forall t \ V(tr_1, Y, t) \leq V(tr_2, Y, t)]$
For all traces tr_1 and tr_2 when in trace tr_1 the levels of state X and Y are always at most as high as the levels of state X and Y in trace tr_2 , for all time points t the value of weight $\omega_{X,Y}$ of the connection from X to Y at t in trace tr_1 is at most the value of weight $\omega_{X,Y}$ at t in trace tr_2	$\forall tr_1, tr_2$ $[\forall t \ [V(tr_1, X, t) \leq V(tr_2, X, t) \wedge V(tr_1, Y, t) \leq V(tr_2, Y, t)] \rightarrow \forall t \ \omega_{X,Y}(tr_1, t) \leq \omega_{X,Y}(tr_2, t)]$
Trace selection property These properties express, for example, that for a given trace the value of a certain state at some time point is at least (or at most) equal to the corresponding value in all other traces. Such trace selection properties are relevant, for example, to find settings of the model that fulfil some property most optimally, for example, the trace with the most significant activation level for a certain state	
For a given trace tr_1 , for all traces tr_2 and all time points t the level of state X at t in trace tr_1 is at least the level of X at t in trace tr_2	$\forall tr_2, t \ V(tr_1, X, t) \geq V(tr_2, X, t)$

This can be expressed, for example by: ‘for any two traces tr_1 and tr_2 if at all time points t in trace tr_2 it holds $V(tr_1, X, t) < V(tr_2, X, t)$ (in trace tr_2 more support X occurs than in trace tr_1), then at all time points t it holds $V(tr_1, Y, t) > V(tr_2, Y, t)$

(in trace tr_2 less stress Y occurs than in trace tr_1)'. For Hebbian learning the following trace comparison property can be formulated:

- *Higher mental state levels imply higher strengths of their connections:*
If in one trace the levels of two connected state are always at least as high as in another trace, then the strength of their connection will always be at least as high in the former trace as in the latter trace.

Trace selection properties

Trace selection properties express that one given trace is special in that it satisfies some property the best within a set of traces. For example, in the sports context, the race in which somebody realised a distance that is a world record has the property:

- *Identifying a record*
For all athletes the distance in all other trace up to now, is at most as far as the distance in this race.

Such a property can be expressed, for example, for the given trace tr_1 and some (end) time point t by (where X represents the distance):

for any trace tr_2 at time point t in trace tr_2 it holds $V(tr_2, X, t) \leq V(tr_1, X, t)$.

A trace selection property is useful, for example, to find settings (i.e., initial values and parameter values) of the model that fulfil some property most optimally, for example, the settings for a trace with the most significant (highest) activation level for a certain state. The property can be checked successively for different choices of trace tr_1 until a (or the) specific trace tr_1 is found satisfying this property. The property to be optimised can be of different types, and also involve, for example, multiple states and time points, as long as some value occurs in it. Multiple states can be optimised simultaneously, by considering a combined value, for example, the sums of the activation levels of a given set of states for a given set of time points, or the sum of squares of the differences (deviations) of the activation levels of a set of states in comparison to empirically known values for a number of time points. In the latter case, this can be used as a method for parameter tuning based on exhaustive search. For a summary of state comparison and trace comparison properties, see Table 13.5.

13.8 Examples of Dynamic Properties in Some Case Studies

In this section some more examples of dynamic properties are discussed.

Dynamic properties for social response patterns

The first types of dynamic properties addressed are those in relation to the model for social response patterns addressed in Chap. 9. By automated verification they have been checked for generated simulation traces, allowing evaluation of the patterns

for a variety of parameter values. Below the dynamic properties are first introduced in an informally expressed manner, after which their formalisations are shown.

The first property expresses that when another B is met, showing a certain emotion, then within a certain time a response occurs, which can consist of:

- (1) self feels the same as person B
- (2) this feeling is bodily expressed by self
- (3) it is communicated b self to B that B feels this

Social response occurrence

For all time points t_1 , when for all time points $t_2 \geq t_1$ person $B \neq \text{self}$ is present and expresses a certain feeling b at t_2 , there exists a time point $t_3 \geq t_1$ such that at t_3 self will have a response r_b (generating the feeling of b , resp. bodily expression, resp. communication with at least a certain level).

$$\forall t_1 [\forall t_2 \geq t_1 [\text{ws}_B(t_2) \geq M_1 \wedge \text{ws}_{B,b}(t_2) \geq M_1] \rightarrow \exists t_3 \geq t_1 r_b(t_3) \geq M_2]$$

with $r_b(t_3)$ one of $\text{srs}_b(t_3)$, $\text{es}_b(t_3)$, $\text{esc}_{B,b}(t_3)$

In case of regulation due to enhanced sensory sensitivity (see Chap. 9), patterns occur when a response only lasts for a short time. This means that the response is withdrawn, which is expressed as follows:

Social response withdrawal

For all time points t_1 and $t_3 \geq t_1$,

when

at t_3 the person self has response r_b with level $\geq M_2$

and

for all

time points $t_2 \geq t_1$

at t_2 person $B \neq \text{self}$ is present with level $\geq M_1$

and

at t_2 expresses a certain feeling b with level $\geq M_1$,

there exists

a time point $t_4 \geq t_3$ such that

$$t_4 \leq t_3 + D$$

and

at t_4 the response r_b has level $< M_3$.

$$\forall t_1, t_3 \geq t_1 [r_b(t_3) \geq M_2 \wedge \forall t_2 \geq t_1 [\text{ws}_B(t_2) \geq M_1 \wedge \text{ws}_{B,b}(t_2) \geq M_1]] \rightarrow$$

$$\exists t_4 \geq t_3 [t_4 \leq t_3 + D \ \& \ r_b(t_4) < M_3]$$

with $r_b(t_3)$ one of $\text{srs}_b(t_3)$, $\text{es}_b(t_3)$, $\text{esc}_{B,b}(t_3)$

The combination of the above two properties specifies a short occurrence of a response. However, after withdrawal of the response due to regulation, also the arousal level for b will become low, which brings back practically the same state as initially. The whole process of responding can start again. An oscillatory pattern results, while the environment is fully static. This is an example of a *limit cycle* pattern. Such oscillatory social responsesocially aware patterns indeed can be observed in persons with some forms of ASD, who let their gaze go back and forth to another person's eyes during a contact, as a way of regulation of enhanced sensitivity. Such a response pattern can be expressed as follows.

Social response oscillation

When B bodily expressing a certain feeling is present from some point in time on, then:

- (1) for every time point there is a later time point for which response r_b occurs
- (2) for every time point there is a later time point for which response r_b does not occur

$$\begin{aligned} \forall t_1 [[\forall t_2 \geq t_1 [ws_B(t_2) \geq M_1 \wedge ws_{B,b}(t_2) \geq M_1] \rightarrow \\ \forall t_3 \geq t_1 [\exists t_4 \geq t_3 [r_b(t_4) > M_2] \& \exists t_5 \geq t_3 [r_b(t_5) < M_3]]] \\ \text{with } r_b(t_3) \text{ one of } srs_b(t_3), es_b(t_3), esc_{B,b}(t_3) \end{aligned}$$

Note that more than these three states are fluctuating. In a more strict quantitative form, a pattern of response oscillation can also be described by the following limit cycle property.

Social response limit cycle

There is a time duration D such that for each of the states r_b from srs_b , es_b and $esc_{B,b}$ for each time point t the value of r_b at $t + D$ is the same as the value of r_b at t .

$$\exists D \forall t \ srs_b(t + D) = srs_b(t) \wedge es_b(t + D) = es_b(t) \wedge esc_b(t + D) = esc_b(t)$$

Dynamic properties for desires and their representation relations

The second types of dynamic properties discussed relate to desires. Desires are assumed to drive behavior in order to achieve 'something that is desired'. A well-known agent model is the BDI model in which desires and beliefs together determine the intentions and actions of an agent. This suggests forward representation relations to the actions that are performed. But what is not addressed in the BDI model is how desires occur. In general this is considered a hard question, maybe partly due to the fact that different types of desire exist, with different types of triggers, for example, a detected unbalance in body state (e.g., hungriness) or observed food as a stimulus.

As in general can be done for cognitive states, desires can be considered not only to represent certain states in the past (*backward representation property*), but also states in the future (*forward representation property*). Looking backward in time, for some cases a desire can be viewed as a mental state representing a certain

unbalance that has developed in a person. For at least some types of desires it is believed that they relate to body states and unbalances in it such as a low blood sugar level. A person needs to generate desires based on represented information about the current body state. The desires generated based on such a body representation tune behaviour to the actual situation of the body. The idea is that the levels of certain body states have to be maintained between certain boundaries (*homeostasis*), for example, the level of the blood sugar state indicated by ws_{bs} . If such a level becomes too low a desire is generated for an action to remedy this, for example, a desire state ds_{eat} to eat something. Then a backward representation relation can be defined based on this.

How a desire represents a (past) body state: backward representation relation

If at some point in time t_1 in a person's body state has a low level of blood sugar, then at a later time point t_2 within time duration D after t_1 this person will have a desire for eating.

If at some point in time t_2 in a person has a desire for eating, then at an earlier time point t_1 within time duration D before t_2 this person's body state had a low level of blood sugar

$$\begin{aligned}\forall t_1 ws_{bs}(t_1) \leq 0.4 &\rightarrow \exists t_2 \geq t_1 [t_2 \leq t_1 + D \wedge ds_{eat}(t_2) \geq 0.7] \\ \forall t_2 ds_{eat}(t_2) \geq 0.7 &\rightarrow \exists t_1 \leq t_2 [t_1 \geq t_2 - D \wedge ws_{bs}(t_1) \leq 0.4]\end{aligned}$$

This property formulation shows how desires emerge from (unbalances in) body states. Note that D gives a time constraint for the time points. The precise value of this parameter can be set to get a realistic property.

However, as desires drive and focus behaviour, a given desire or combination of desires can also be considered while looking forward in time. From that perspective the occurrence of a desire at some point in time can be considered to represent the future performance of a certain action.

The following dynamic property can be considered such a forward representation property for a desire. It expresses that when desires for a certain type of action occurs, then this type of action will be performed.

How a desire represents a (future) action: forward representation relation

If at some point in time t_1 a person desires an action of some type, then at a time point t_2 within time duration D after t_1 this person will perform an action of that type (and conversely).

$$\begin{aligned}\forall t_1 ds_{eat}(t_1) \geq 0.7 &\rightarrow \exists t_2 \geq t_1 [t_2 \leq t_1 + D \wedge es_{eat}(t_2) \geq 0.7] \\ \forall t_2 es_{eat}(t_2) \geq 0.7 &\rightarrow \exists t_1 \leq t_2 [t_1 \geq t_2 - D \wedge ds_{eat}(t_1) \leq 0.7]\end{aligned}$$

In this formulation the property shows that the indicated desire drives the person to a specific type of action, and such an action is grounded in a desire preceding it.

A forward representation property for desire describes, for example, the behaviours or actions that are driven or motivated by these desires. For the behaviour of

a person to make sense, these actions are not arbitrary, but are usually assumed to achieve fulfillment of the desires. So, looking forward in time some steps further, desires can be viewed as indicating a future state in which they are fulfilled. As after a desire has been fulfilled, in principle it is not there anymore, such a forward relation may sound a bit circular and paradoxal: desires that exist in the present indicate their own future nonexistence. The following example of a dynamic property for fulfillment of a desire; it can be viewed as a specific achievement property:

Relating a desire to its fulfilment

If at some point in time t_1 a person desires a certain action a , then at a time point t_2 within time duration D after t_1 the person will not desire this anymore.

$$\forall t_1 \text{ ds}_a(t_1) \geq 0.7 \rightarrow \exists t_2 \geq t_1 [t_2 \leq t_1 + D \wedge \text{ds}_a(t_2) \leq 0.3]$$

13.9 Automatic Checking of Dynamic Properties

A software environment has been developed to automate the checking process for dynamic properties, based on the Temporal Trace Language TTL; see (Bosse et al. 2009a; Sharpanskykh and Treur 2010). Such a software environment takes as input a dynamic property specification and a trace (or set of traces) and returns as an outcome whether this property is true or not in this trace (or set of traces). For an implementation it is convenient to use a logical language such as PROLOG. But as logical languages often require a slightly different format, first some transformations have to be applied to the dynamic property, in particular, to eliminate the use of functional expressions such as $V(tr, X, t)$.

Transformations of dynamic properties

To express dynamic properties in implemented logical languages such as TTL, or PROLOG, they may need some standard transformations. Often in such logical languages basic, functional expressions involving numbers such as $V(tr, X, t_1)$ have to be represented in a different, relational manner. A standard way is to replace them by variables V_i for the indicated values, relate this value to the name of the state and use a quantifier for this value; for example using existential quantifiers:

$$\begin{aligned} V(tr, X, t_1) < V(tr, X, t_2) : \\ \exists V_1, V_2 [\text{has_value_at}(tr, X, t_1, V_1) \wedge \text{has_value_at}(tr, X, t_2, V_2) \wedge V_1 < V_2] \end{aligned}$$

Another option using universal quantifiers:

$$\begin{aligned} V(tr, X, t_1) < V(tr, X, t_2) : \\ \forall V_1, V_2 [\text{has_value_at}(tr, X, t_1, V_1) \wedge \text{has_value_at}(tr, X, t_2, V_2) \rightarrow V_1 < V_2] \end{aligned}$$

The reified temporal predicate logical language TTL

The language TTL (Jonker and Treur 2002; Bosse et al. 2009b; Sharpanskykh and Treur 2010) is an example of a reified temporal predicate logic language; e.g., (Ma and Knight 2001; Galton 2006). This means in particular that states X and state properties P can be represented as individuals which can also be indicated by variables over which quantifiers can be used, and the same holds for traces tr . Here state properties P are specific properties of an overall state such as

$\text{has_value}(X, V)$

expressing that state X has value V , or

$\text{has_value}(X_1, V_1) \wedge \text{has_value}(X_2, V_2) \wedge V_1 < V_2$

expressing that state X_1 has value V_1 and state X_2 has value V_2 and $V_1 < V_2$.

Within TTL the predicate

$\text{holds}(S, P, TV)$

relates an overall state S to a state property P in such a manner that truth value TV is true indicates that P holds in overall state S and TV is false indicates that it does not hold in overall state S . Moreover, for a given trace tr to indicate the overall state in a trace tr at some time point t the expression

$\text{state}(tr, t)$

is used. So, for example,

$\text{holds}(\text{state}(tr, t), P, \text{false})$

indicates that within the overall state of trace tr at time t state property P is false.

To obtain expressions in TTL the following substitutions can be done:

$$\begin{aligned} V(tr, X, t_1) < V(tr, X, t_2) : \\ & \exists V_1, V_2 \\ & [\text{holds}(\text{state}(tr, t_1), \text{has_value}(X, V_1), \text{true}) \wedge \\ & \text{holds}(\text{state}(tr, t_2), \text{has_value}(Y, V_2), \text{true}) \wedge V_1 < V_2] \end{aligned}$$

Or using universal quantifiers:

$$\begin{aligned} V(tr, X, t_1) < V(tr, X, t_2) : \\ & \forall V_1, V_2 \\ & [\text{holds}(\text{state}(tr, t_1), \text{has_value}(X, V_1), \text{true}) \wedge \\ & \text{holds}(\text{state}(tr, t_2), \text{has_value}(Y, V_2), \text{true}) \rightarrow \\ & V_1 < V_2] \end{aligned}$$

In this way the reified temporal predicate logic language TTL can be used to express the wide variety of dynamic properties in a formal format, including properties in which states and traces are explicitly represented and compared. This makes TTL a more expressive language than most other temporal languages; e.g., see (Bosse et al. 2009a; Sharpanskykh and Treur 2010). In particular, all example properties discussed in this chapter can be expressed in TTL.

Using the TTL environment

A TTL software environment is available with an editor to specify dynamic properties but also with a checker to automatically check such dynamic properties against given traces (or sets of traces). This can be used for the different types of applications described in Sects. 13.3 and 13.4.

Monitoring and analysis of dynamic properties in real world processes

The traces used as input for the TTL software environment can be based on real world data, for example, obtained by monitoring an ongoing process. Within the TTL environment, as a form of analysis of these monitoring data it can be checked on the fly whether some complex dynamic properties hold for these ongoing processes. In this way that the TTL environment can play a role in monitoring and analysis of complex dynamic patterns in real world contexts.

Extensive testing, analysis and personalisation of a model

Using a simulation model a large variety of traces can be generated, for example, to explore initial values, input from the environment and parameter values by systematically or randomly varying them. If these generated traces are fed into the TTL environment it will automatically be determined which of these traces have some specified dynamic properties. As all of this can be automated, this can support extensive forms of testing, analysis and tuning of a model, and, for example, be used to find out how the emergence of some dynamic property depends on the settings (initial values, external input from the environment over time, and parameter values), which can be used to obtain some extent of predictability of emergent properties depending on such settings.

13.10 Discussion

In this chapter different types of dynamic properties describing patterns in behaviour dynamics that can emerge from the local mechanisms described by the connections in a temporal-causal network have been discussed. It has been shown how these patterns can be expressed both by a conceptual representation and a (formal) numerical-logical representation, and how they can be systematically checked for a given trace describing the states of a process over time. Moreover, it has been discussed how such checking can be automated by using a dedicated software environment such as the TTL environment described in (Bosse et al. 2009b; Sharpanskykh and Treur 2010), and how in particular the dynamic

properties can be validated in an automated manner when real world data on process traces are available.

It has been discussed how dynamic properties can be useful when they are applied, both in the context of monitoring and analysis of a real world process either off-line or on the fly, and in the context of extensive testing and analysis of a network model as a form of focusing and verification of the model during Network-Oriented Modeling. In the latter case, large numbers of traces generated by the model can be analyzed automatically. It has been discussed how as a result of this analysis, emergent properties of such a model can become predictable (to a certain extent) from the model and its settings (initial values, external input from the environment, parameter values). Finally, it was pointed out how the perspective discussed here can be used to validate a dynamical model in an automated manner, and to find values of parameters to satisfy required properties of the model and in this way personalize the model.

Temporal perspectives on properties of the dynamics of processes are addressed in a wide variety of literature, from different disciplines, including Cognitive Science, Artificial Intelligence and Computer Science. Literature on dynamics and temporal property specification and verification from an AI or Computer Science perspective can be found in, for example, (Barringer et al. 1996; Bosse et al. 2009b; Clarke et al. 2000; Goldblatt 1992; Juan and Tsai 2002; Manna and Pnueli 1990, 1995; Stirling 2001). In this literature different types of properties are distinguished, of the types which were also discussed in this chapter. However, usually in these approaches it is not possible to express state comparison, trace comparison and trace selection properties, as the languages used do not support the reification that is needed in order to use quantifiers over states and traces for these types of properties. Moreover, most of these languages are purely logical and have less adequate numerical expression possibilities.

References

- W.R. Ashby, *Design for a Brain* (Chapman & Hall, London. Revised edition, 1960)
- H. Barringer, M. Fisher, D. Gabbay, R. Owens, M. Reynolds, *The Imperative Future: Principles of Executable Temporal Logic* (Research Studies Press Ltd. and Wiley, 1996)
- T. Bosse, C.M. Jonker, L. van der Meij, A. Sharpanskykh, J. Treur, Specification and verification of dynamics in agent models. *Int. J. Cooperat. Inf. Syst.* **18**, 167–193 (2009a)
- T. Bosse, C.M. Jonker, J. Treur, Representation for reciprocal agent-environment interaction. *Cogn. Syst. Res.* **10**, 366–376 (2009b)
- E.M. Clarke, O. Grumberg, D.A. Peled, *Model Checking* (MIT Press, 2000)
- J. Ferber, O. Gutknecht, C.M. Jonker, J.P. Müller, J. Treur, Organization models and behavioural requirements specification for multi-agent systems, in *Multi-Agent System Organisations. Proceedings of the 10th European Workshop on Modelling Autonomous Agents in a Multi-Agent World, MAAMAW'01*, ed. by Y. Demazeau, F. Garijo (2001)
- A. Galton, Operators vs. arguments: the ins and outs of reification. *Synthese* **150**, 415–441 (2006)
- R. Goldblatt, *Logics of Time and Computation*, 2nd edn. (CSLI Lecture Notes 7, 1992)

- D.E. Herlea, C.M. Jonker, J. Treur, N.J.E. Wijngaards, Specification of behavioural requirements within compositional multi-agent system design, in *Multi-Agent System Engineering, Proceedings of the 9th European Workshop on Modelling Autonomous Agents in a Multi-Agent World, MAAMAW'99. Lecture Notes in AI*, vol. 1647, ed. by F.J. Garijo, M. Boman (Springer, Berlin, 1999), pp. 8–27
- D.E. Herlea Damian, C.M. Jonker, J. Treur, N.J.E. Wijngaards, Integration of behavioural requirements specification within compositional knowledge engineering. *Knowle. Based Syst. J.* **18**, 353–365 (2005)
- C.M. Jonker, J. Treur, Compositional verification of multi-agent systems: a formal analysis of pro-activeness and reactiveness. *Int. J. Cooper. Inf. Syst.* **11**, 51–92 (2002)
- E.Y.T. Juan, J.J.P. Tsai, *Compositional Verification of Concurrent and Real-Time Systems* (Kluwer Academic Publishers, 2002)
- J. Kim, *Philosophy of Mind* (Westview Press, 1996)
- A. van Lamsweerde, *Requirements Engineering: From System Goals to UML Models to Software Specifications* (Wiley, Chichester, 2009)
- J. Ma, B. Knight, Reified temporal logics: an overview. *Artif. Intell. Rev.* (2001)
- Z. Manna, A. Pnueli, A hierarchy of temporal properties, in *Proceedings of the Ninth Annual ACM Symposium on Principles of Distributed Computing, PODC'90* (ACM Press, 1990), pp. 377–410
- Z. Manna, A. Pnueli, *Temporal Verification of Reactive Systems: Safety* (Springer, 1995)
- K. Pohl, *Requirements Engineering—Fundamentals, Principles, and Techniques*, 1st edn. (Springer, 2010)
- K. Pohl, C. Rupp, *Requirements Engineering Fundamentals* (Rocky Nook, 2011)
- R.F. Port, T. van Gelder (eds.), *Mind as Motion: Explorations in the Dynamics of Cognition* (MIT Press, Cambridge, Mass, 1995)
- A. Sharpanskykh, J. Treur, A temporal trace language for formal modelling and analysis of agent systems, in *Specification and Verification of Multi-Agent Systems*, ed. by M. Dastani, K.V. Hindriks, J.J.C. Meyer (Springer, 2010), pp. 317–352
- C. Stirling, *Modal and Temporal Properties of Processes*, vol. 15 (Springer, 2001), pp. 189–217
- D.J. Thilakarathne, A parameter estimation method for dynamic computational cognitive models, in *Proceedings of the 6th Annual International Conference on Biologically Inspired Cognitive Architectures, BICA'15, Procedia Computer Science*, vol. 71 (2015), pp. 133–142

Chapter 14

Who are You

Identifying Characteristics of Persons, Their Networks and Other Contextual Aspects by Parameter Estimation and Validation

Abstract In this chapter it is discussed how a personalised temporal-causal network model can be obtained that fits well to specific characteristics of a person, and his or her connections and further context. A model is a close approximation, but always a form of abstraction of a real world phenomenon. Its accuracy and correctness mainly depend on the chosen abstracting assumptions and the values of the parameters in the model representing contextual characteristics. Depending on the complexity of the model, the number of its parameters can vary from just a couple to thousands. These parameters usually represent specific characteristics of the modeled phenomenon, for example, for modeling human processes personality characteristics or social interaction characteristics. No values for such parameters are given at forehand. Estimation of parameters for a given model is a nontrivial task. There are many parameter estimation methods available in the literature. In this chapter a number of these methods are briefly discussed.

14.1 Introduction

Dynamical models such as temporal-causal network models usually have to take into account a number of characteristics of the situation that is modeled. Such characteristics can involve, for example, the mental or neurological structures of a person, or a person's connections to others, or contextual elements of the external world. Usually in a model description parameters are used to represent such characteristics. The advantage of having such parameters in a model description is that they enable to use and tune the model for different situations: for example, persons with different mental or neurological structures, for different social connections, or for different contextual elements in the external world. In fact the model represents a large (and in theory infinite) space of possibilities indicated by all combinations of values of the parameters. For example, suppose 10 parameters are involved and all parameters are in the interval $[0, 1]$. If only parameter values in one digit are considered (i.e., 0, 0.1, ..., 0.9, 1.0), then the number of combinations already is 11^{10} , which is more than 25 billion or 2.5×10^{10} . If the parameter values

are considered in two digits, this number will be more than 100^{10} , which is 10^{20} . So, a model with a number of parameters is very generic in the sense that the space of situations that can be represented by the model can be huge, with many variations.

For one given specific situation at hand the parameters have to be assigned values that represent that situation in particular; the parameter values have to be found that fit to the situation: by finding such values, knowledge of the specific characteristics of the situation is acquired. However, such a tuning to specific characteristics is not always easy, as often a situation that is modeled does not simply show these characteristics. They have to be acquired or estimated by some process in one way or the other, and this may turn out not so easy. In Fig. 14.1 such an estimation process is sketched. The observed behaviour from the actual phenomenon is compared to the predicted behaviour from the dynamical model at some time points. If there is a significant difference (the objective should be to make this difference minimal), the model's behaviour has to be made closer to the observed behaviour by changing the values of the parameters. More specifically, the various parameter estimation methods specify how to quantify the difference and based on that in what way the values of the parameters should be adjusted.

In the current chapter this issue is discussed in some detail. In Sect. 14.2 it is discussed how in the many cases that direct measuring is hard or impossible, via requirements the values of parameters can be determined. In Sect. 14.3 an example of a temporal-causal network model is shown that is used in this chapter to illustrate the different methods. In Sects. 14.4–14.7 four different approaches are discussed: exhaustive search, gradient descent, random gradient descent, and simulated annealing, respectively.

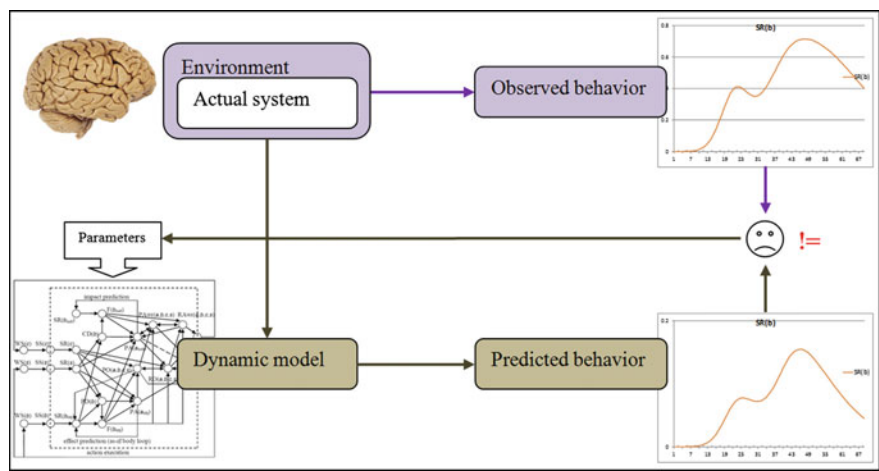


Fig. 14.1 Parameter estimation for a dynamical model

14.2 Determining Characteristics and the Use of Requirements

Without having precise knowledge of the specific characteristics of a situation that is modeled, it will be difficult to obtain a model that really fits to the situation. In a model usually a number of parameters are used to represent such characteristics. In Sect. 14.2.1 the specific parameters in a temporal-causal network model are discussed. The problem to find proper values for such parameters representing characteristics for a given situation basically can be addressed in two manners: by direct measuring of the characteristics in the situation that is modeled, or via requirements. These will be discussed in some more detail in Sects. 14.2.2 and 14.2.3.

14.2.1 *The Parameters in a Temporal-Causal Network Model*

Within a temporal-causal network model, in particular the following types of parameters occur:

- Connection weights $\omega_{X,Y}$ for states X and Y
- Speed factors η_X for each state X
- Parameters within specific combination functions, such as:
 - in a scaled sum combination function the scaling factor λ
 - in a logistic sum combination function the parameters σ and τ
- For models with adaptive connections, for example:
 - for Hebbian learning the learning rate η and the extinction rate ζ
 - for adaptive networks based on the homophily principle the threshold $\tau_{X,Y}$ and speed factor $\eta_{X,Y}$ for states X and Y

Here, for example, in a specific situation the connection weights may relate to the strengths of certain connections in someone's brain or to the strengths of certain connections in social interaction, and speed factors may relate to actual speed of processing the states. For a given situation it is not clear at forehand how values of such parameters have to be chosen. There are some indications or heuristics that can be kept in mind, to manage them during modeling:

- Connection weights 1 for maximal effect and lower between 0 and 1 for a smaller effect
- Connection weights between -1 and 0 for suppressing effects
- Speed factors can be chosen higher for internal, mental processes, and lower for body changes and execution of actions in the world
- For scaled sum combination function: choose the scaling factor λ equal to or at most the sum of the weights of the incoming connections

- For logistic sum combination functions
 - Choose steepness σ low (from 2 to 8) for gradual effects and high (10–20 or higher) for more all or nothing types of effects
 - The threshold τ usually has a relation with the number of incoming connections and their weights. For example, the aggregated impact can never be more than the sum of these weights, so a threshold higher than that will not lead to substantial activation of a state; it usually is higher if there are more incoming connections

Such heuristics still will not make it easy to find values that adequately represent the specific characteristics of a given situation. This basically can be addressed in two manners: by direct measuring of the characteristics in the situation that is modeled, or via requirements. These will be discussed in some more detail in Sects. 14.2.2 and 14.2.3.

14.2.2 *Direct Measuring of Characteristics of a Situation*

From a naïve point of view, the possibility to directly measure values of parameters is the most attractive option. For example, if some physical process is to be modeled, according to some physical laws in which certain quantities (such as mass or volume) occur as parameters, then the values of these quantities can be measured and used for these parameters. This may work in an idealised physical domain, but for human and social domains this may be less straightforward. Suppose a connection from one mental state X to another mental state Y is involved in the model, then according to the current state of the art measuring the strength of this connection is quite difficult, if not impossible.

As another example, suppose in a network that models social interaction, a connection from person X to person Y occurs. How could the strength of this connection be measured? By the number of Whatsapp messages per minute? By the time duration of telephone calls? By the time duration of being at the same location? As discussed in Chap. 11, Sect. 11.9, in the literature it is discussed how connection strength in networks describing social interaction relates to aspects such as interaction frequency, emotional intensity of content, and emotional support and closeness (Gilbert and Karahalios 2009; Granovetter 1983; Marsden and Campbell 1990). In part of this literature the relation between connection strength and aspects of actual interaction is used to formulate a measurable definition of connection strength. However, in other literature not a definitional but a causal relationship between such measurable aspects and connection strength is assumed; e.g., Hove and Risen (2009), Pearce et al. (2015). So also direct measuring of connection strength in a network describing social interaction is not without problems.

As another alternative, sometimes questionnaires are used for measurements of characteristics in human or social domains, where persons can score their

characteristics. At first sight this may seem practical and adequate, as such a scoring process generates numbers, and for parameters numbers are needed. But this also has some problems. First of all, persons do not necessarily know their own characteristics, and if they believe they do know, there is no guarantee that these beliefs are correct. And secondly, a score from a questionnaire, or an aggregation of such scores, may provide a number, but this number is supposed to be measured according to some scale, and it may not be clear how this scale relates to the scale of a relevant parameter. There may be a nontrivial, unknown relation between such scales, perhaps at least a monotonic relation, but maybe not proportional or linear. So to adequately translate such scores into values for parameters can be a problem by itself.

14.2.3 Using Requirements to Find Characteristics of a Situation

Another way to tune parameters is to identify and explicate what the model is expected to do: expressed as requirements for the model; see also Chap. 13. Suppose such requirements have been identified, and it has been found that they indeed describe what is expected from the model. Then a number of values of parameters can be tried alternatively until values are found such that the model shows the behaviour fulfilling the requirements. Usually this is already done intuitively by a modeler. However, for larger numbers of parameters the huge space of possibilities for a model now turns into a huge search space. For models with many parameters it is easy to get lost in the large search space of all combinations of parameter values. So, eventually the question how to find proper values for the parameters may get an answer in the form of a search problem that is to be solved. The requirements used for this search problem can be of different types. In general the requirements can take the form of any temporal patterns expressed as dynamic properties as addressed in Chap. 13. Different combinations of parameter values can be tried in order to find those combinations of parameter values that lead to fulfilment of the requirements. This process can be performed in a systematic manner, as exhaustive search, or some form of heuristic search. An example of such a heuristic process for a temporal-causal network model can be found in Thilakarathne (2015).

For heuristic search usually some measure is used to indicate how far from fulfilment the requirement is; this is often called an *error measure* or *error function*. The requirements can be of a very specific form when for some states of the model *empirical values* are available for some of the time points, and it is considered that the model is required to generate values at these time points equal or close to the empirical values.

14.2.4 Using Error Measures for Requirements

The heuristic methods discussed below in Sects. 14.5–14.7 assume an error measure. In many cases parameter estimation methods are applied when a data set is available for a given (empirical) trace represented by values for certain states and time points, so basically a set of triples (state, time, value). The requirement considered is that a trace generated by the model shows values for the states at these time points that are equal to the values indicated by the given empirical trace, or at least close to these values. As being equal is usually not feasible, the question becomes how to define this ‘being close to’ for multiple states and time points.

An error function expresses in one way or the other in an aggregated manner the deviation of the simulation values for the considered states and time points in comparison to the empirically given values. When there is no deviation for any of the states and time points, the error function will give the value 0, and if the deviation is small, the value of the error function will be close to 0. For a parameter estimation method the aim is to get the value of the error function below some small value (accuracy) or as close to 0 as possible.

Error functions usually take the differences between simulated values and empirical values for all given data points (state, time, value) as a point of departure. Suppose for some (but usually not all) states X_i , $i = 1, 2, \dots$, for some (but maybe not all) time points $t_{i,j}$, $j = 1, 2, \dots$ empirical values $\underline{V}_{i,j}$, $j = 1, 2, \dots$ are given. This forms an empirical data set $(\underline{X}_i, t_{i,j}, \underline{V}_{i,j})$, $i = 1, 2, \dots, j = 1, 2, \dots$. For these time points $t_{i,j}$ the simulated values for a specific simulation trace $\underline{tr}$ (according to some scenario) are indicated by $V(\underline{tr}, \underline{X}_i, t_{i,j})$. Then the relevant differences (also called deviations or residuals) are

$$D_{i,j} = V(\underline{tr}, \underline{X}_i, t_{i,j}) - \underline{V}_{i,j}$$

The strict way of expressing such a requirement is to state that these differences are all 0:

$$V(\underline{tr}, \underline{X}_i, t_{i,j}) = \underline{V}_{i,j}$$

for all states X_i and time points $t_{i,j}$. Sometimes this may be relevant and feasible, but often this is difficult to achieve. A less strict and more often applicable way of expressing the requirement, for example, that all the absolute values of the differences should be at most a given small positive number D :

$$|D_{i,j}| \leq D$$

for all states X_i and time points $t_{i,j}$. Here D is a small positive number, for example 0.05. Yet another way is to aggregate the deviations $D_{i,j}$ into one number, which is called an error value or error function. One possible way of aggregation is to take the average deviation for all triples in the data set, but this is not satisfactory as the

positive and negative differences cancel each other out. A better option is to take the average absolute deviation:

$$\text{error} = (|D_{1,1}| + |D_{1,2}| + \cdots + |D_{2,1}| + |D_{2,2}| + \cdots + \dots)/N$$

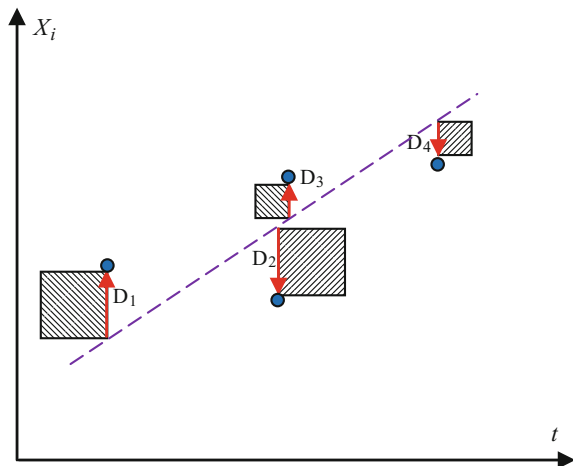
or in short notation: $\sum_{i,j} |D_{i,j}|/N$. Here N is the total number of data points used. However, an often used form of aggregation for an error function makes use of the sum of squares of the differences $D_{i,j}$ (for a visualisation, see Fig. 14.2): $\sum_{i,j} D_{i,j}^2$.

The sum of squares of residuals as a basis for an error function is a generic concept that is used in applications in many disciplines to measure the deviation of a set of data points relative to a reference; the reference can be a (partial) curve described or generated by a mathematical model or some set of known data values. The word residual refers to the difference between observed vs predicted values for the considered variable. Minimizing the sum of squares of residual values is referred as a *least square method* (Moler 2004). The history of the least square method goes back to 1795, when Karl Friedrich Gauss has formulated it as a basic concept and found out that when it is assumed that measurements deviate from an ideal pattern according to a normal probability distribution, then a least square method provides an optimal approximation of the ideal pattern; see, for example, Strejc (1980).

To make the values obtained by calculating the sum of squares comparable with the actual differences $E_{i,j}$ it is useful to apply the square root of the average of the squares:

$$\text{error} = \sqrt{\sum_{i,j} D_{i,j}^2 / N}$$

Fig. 14.2 Sum of squares: graphical representation of the areas of the squares



This makes it much easier and intuitive to verify whether the error makes sense in comparison to interpreting the overall areas of the squares in terms of vertical distances between points in the plane. The values of such an error function directly relate to the (linear) vertical differences that can be seen in the graph. For example, when $N = 4$ and all deviations $D_{i,j}$ are the same D , then this results in error D and not in $4D^2$, as the sum of the squares would do, and what is difficult to relate to the vertical distances in a graph as shown in Fig. 14.2.

In this way an error function based on the sum of squares of residuals can be used in a practical manner to evaluate the quality of the selected values of the parameters in a given model in comparison to empirical data. Using any error function there are different ways how to formulate a requirement. One most strict requirement would be that the error is 0. Although sometimes this may be relevant, in many practical situations such a requirement is too strict. Another option is to express in a requirement that the error is at most a given small value D , for example, 0.05:

$$\sqrt{\sum_{i,j} D_{i,j}^2 / N} \leq D$$

This can be used in a generate-and-test method that works by generating traces under systematic variations of the settings of a model one by one and for each trace testing whether this requirement is fulfilled, until one is found that fulfills the requirement.

When a model has many parameters it may be difficult to generate a set of traces tr_j for the many relevant variations of settings for the parameters. In such cases within the process of generation of traces often heuristic search methods are applied, based on what has been found up to some point in time. Such methods will be discussed in Sects. 14.4 and further.

14.3 Description of an Example Model

In each section below a method for parameter estimation is discussed. In these sections, to demonstrate the method a common example of a temporal-causal network model will be used for illustration. This model was inspired by the idea of Damasio's as-if body loop (Damasio 1999; Damasio et al. 2000; Parvizi et al. 2006). This computational model describes the generation of feelings. A conceptual representation of the model is shown in Fig. 14.3 below; the states used in the model are summarized in Table 14.1.

The model uses two inputs: stimulus s , and body state b , which may occur as a response to the stimulus. The stimulus s is associated to an emotional response b leading to a detectable body state ws_b (e.g., a face expression). In turn the effect ws_b serves as input by sensing it.

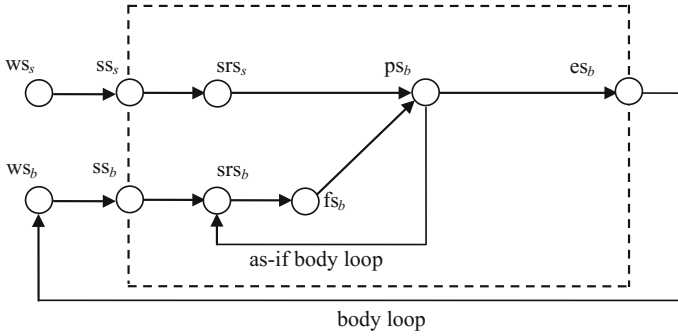


Fig. 14.3 Example model for modeling of feeling

Table 14.1 States in the example model shown in Fig. 14.3

Notation	Description
ws_W	World state for W (W is a stimulus s , or body state b)
ss_W	Sensor state for W
srs_W	Sensory representation state for W
ps_b	Preparation state for response b
fs_b	Feeling state for b
es_b	Execution state for response b

World states ws_w (e.g., ws_s , and ws_b) affect sensor states ss_w (e.g., ss_s , and ss_b respectively). The sensor states lead to further internal processes according to the following causal sequence (the body loop in Fig. 14.3):

sensing a stimulus $ss_s \rightarrow$ sensory representation of stimulus $srs_s \rightarrow$
 preparation for bodily response $ps_b \rightarrow$ execution of action $es_b \rightarrow$
 sensing the bodily response $ss_b \rightarrow$ sensory representation of the bodily response $srs_b \rightarrow$ feeling the emotion fs_b

The effect prediction loop or as-if body loop goes from preparation for bodily response to sensory representation of the bodily response to feeling the associated emotion (Damasio 1999; Damasio et al. 2000). In the effect prediction loop, the preparation for action ps_b is affected by the sensory representation of stimulus srs_s , and the feeling of effect prediction fs_b of action a . The sensory representation srs_b of effect b gets effects from the preparation for action ps_b , and the sensory state ss_b is affected by the execution of response b . The feeling fs_b is affected by the sensory representation srs_b of b .

The connections between state properties (the arrows in Fig. 14.3) have weights ω_k , as indicated in Table 14.2. In this model it is assumed that all weights are non-negative and between 0 and 1.

In the example simulations, for the states Y that are affected by only one state, the combination function $c_Y(\dots)$ is taken as the identity function $c_Y(V) = \text{id}(V) = V$,

Table 14.2 Overview of the connections and their weights

From state	To state	Weight
ws_s	ss_s	ω_1
es_b	ws_b	ω_2
ws_b	ss_b	ω_3
ss_s	srs_s	ω_4
srs_s	ps_a	ω_5
fs_b		ω_6
ss_b	srs_b	ω_7
ps_b		ω_8
srs_b	fs_b	ω_9
ps_b	es_b	ω_{10}

and for the other states ps_b and srs_b the function $c_Y(\dots)$ is a combination function is the advanced logistic sum function **alogistic** $_{\sigma,\tau}(\dots)$.

$$c_Y(V_1, V_2) = \mathbf{alogistic}_{\sigma,\tau}(V_1, V_2) = \left(\frac{1}{1 + e^{-\sigma(V_1 + V_2 - \tau)}} - \frac{1}{1 + e^{\sigma\tau}} \right) (1 + e^{-\sigma\tau})$$

when $V_1 + V_2 > 0$

$$c_Y(V_1, V_2) = 0 \quad \text{when } V_1 + V_2 \leq 0$$

In the latter combination function τ is the threshold and σ is the steepness; for ps_b and srs_b the respective parameters are τ_{ps_b} , σ_{ps_b} , τ_{srs_b} and σ_{srs_b} . For the speed factor η_X two values are used: η_{slow} (slower) for external states X and η_{fast} (faster) for internal states X . As sensor states and execution states need more time to change physically, the speed factor for external states should be low compared to the ones for internal states. In the model the states ws_s , ws_b , ss_s , and ss_b are considered to be external.

Given these the dynamics of the model can be described as follows:

$$\begin{aligned} ss_s(t + \Delta t) &= ss_s(t) + \eta_{\text{slow}}[\omega_1 ws_s(t) - ss_s(t)] \Delta t \\ ws_b(t + \Delta t) &= ws_b(t) + \eta_{\text{slow}}[\omega_2 es_b(t) - ws_b(t)] \Delta t \\ ss_b(t + \Delta t) &= ss_b(t) + \eta_{\text{slow}}[\omega_3 ws_b(t) - ss_b(t)] \Delta t \\ srs_s(t + \Delta t) &= srs_s(t) + \eta_{\text{fast}}[\omega_4 ss_s(t) - srs_s(t)] \Delta t \\ ps_b(t + \Delta t) &= ps_b(t) + \eta_{\text{fast}}[\mathbf{alogistic}_{\sigma_{ps_b}, \tau_{ps_b}}(\omega_5 srs_s(t), \omega_6 fs_b(t)) - ps_b(t)] \Delta t \\ srs_b(t + \Delta t) &= srs_b(t) + \eta_{\text{fast}}[\mathbf{alogistic}_{\sigma_{srs_b}, \tau_{srs_b}}(\omega_5 ss_b(t), \omega_6 ps_b(t)) - srs_b(t)] \Delta t \\ fs_b(t + \Delta t) &= fs_b(t) + \eta_{\text{fast}}[\omega_9 srs_b(t) - fs_b(t)] \Delta t \\ es_b(t + \Delta t) &= es_b(t) + \eta_{\text{slow}}[\omega_{10} ps_b(t) - es_b(t)] \Delta t \end{aligned}$$

For the expected feeling when the stimulus has level 1, the data was chosen as shown in Fig. 14.3. For this case study these were generated by the model, after which according to a normal distribution some noise was added to make them look like empirical data. The parameters of the model are 10 connection weight values

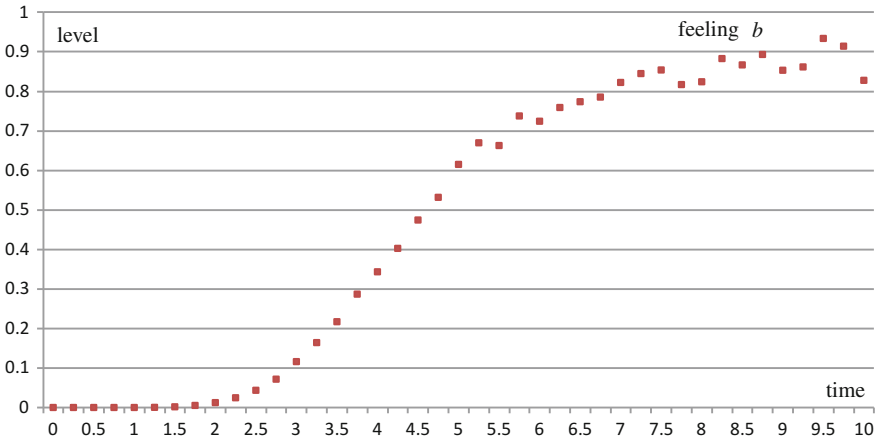


Fig. 14.4 The required behaviour of fs_b when the world state for stimulus ws_s is 1

(ω_1 – ω_{10} in Table 14.2), 2 threshold values (τ_{ps_b} and τ_{srs_b}), 2 steepness values (σ_{ps_b} and σ_{srs_b}), 2 speed factors (η_{slow} and η_{fast}). It is assumed that $\eta_{slow} = 0.5$ and $\eta_{fast} = 0.9$, and that the time step size is $\Delta t = 0.25$. The remaining 14 parameter values are to be determined by parameter estimation.

As an illustration to this example different parameter estimation methods will be applied in subsequent sections. By applying some (relative) noise to the values of an example simulation according to a normal distribution, the data shown in Fig. 14.4 were obtained. By parameter estimation values of parameters will be determined that make that the simulated feeling level approximates these values shown in Fig. 14.4. As error function the square root of the average of the squares of deviations has been used, and as a stop criterion for the estimation process an error of 0.03.

Note that no unique solution can be expected. There may be different types of persons with different characteristics that still generate similar feeling levels. Different parameter estimation methods and different settings and initial values of them in principle will generate different solutions. Parameter estimation can be used to explore the different possibilities.

14.4 Parameter Tuning by Exhaustive Search

Exhaustive search (also called brute-force search) is a quite elementary method for parameter tuning. It consists of

- Systematically enumerating all possible assignments of values to the parameters for the model, with a certain grain size or accuracy, for example in two or three decimals

- For each of these assignments run the model to generate a simulation trace
- For each generated simulation trace check whether (and to which extent) the generated simulation trace fulfils the requirements.

The option(s) that fulfill the requirements (or fulfill them best) are selected as suitable options for values of the parameters. If requirements are not fulfilled in a strict sense but only in an approximate sense, it is assumed that an error function is used, and the parameter values which show the lowest value for this error function can be chosen as the best outcome; here as error function, for example, the square root of the average of the squares of residuals may be used. For this case, example pseudo-code is shown in Box 14.1.

```

AP = first candidate assignment of values % Initialization
least error =  $10^{-10}$  % Initialization
while (no terminate while) % For all possible parameter assignments
  run the model for AP % generate a simulation trace
  if (error(AP) < least-error ) % if this assignment is better than the best one found yet
    best-set-of-parameter-values = AP % the new set of parameters is the best until now
    least-error = error(AP) % the error of the new set is the least one
  end if
  if assignment left
    AP = next-assignment(AP) % go to the next value assignment
  else
    terminate-while
  end if
end while

```

Box 14.1 Pseudo-code for exhaustive search

Usually dynamical models have continuous parameters. To use exhaustive search for such models, the parameter values have to be assigned discrete values. The simplest way to do this is according to a uniform grain size, for example of 0.01; this grain size is a measure for the accuracy by which the search is performed. For example, suppose that the model has just one parameter P_1 , which is continuous and it varies between 0 and 1. The aim is to find a value for P_1 which minimizes the difference between empirical data (observation) and model prediction. For example, for accuracy 0.1, parameter P_1 can be assigned discrete values 0.0, 0.1, 0.2, 0.3, ..., 0.9, 1.0, respectively. According to the exhaustive search method, the error (difference between empirical data and model prediction) for each of these values has to be determined, and the one which leads to the lowest error can be chosen.

As it is clear in Table 14.3, in this example, the error is minimal when P_1 is equal to 0.1. Thus, for this fictitious example, the exhaustive search method suggests to choose the value 0.1 for parameter P .

Table 14.3 Example error values for different parameter values

P ₁	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
Error	0.443	0.170	1.084	2.010	2.731	3.265	3.665	3.972	4.218	4.421	0.443

Table 14.4 Error of the prediction of the model with different set of parameters

P ₁ P ₂	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0.0	0.114	0.083	0.064	0.056	0.061	0.077	0.105	0.145	0.197	0.261	0.337
0.1	0.058	0.022	0.003	0.022	0.02	0.058	0.116	0.193	0.292	0.412	0.553
0.2	0.12	0.086	0.077	0.096	0.144	0.223	0.333	0.476	0.655	0.87	1.124
0.3	0.299	0.278	0.294	0.35	0.449	0.594	0.789	1.037	1.342	1.708	2.139
0.4	0.596	0.603	0.661	0.776	0.953	1.199	1.52	1.923	2.416	3.008	3.706
0.5	1.011	1.062	1.184	1.385	1.675	2.065	2.566	3.191	3.954	4.871	5.958
0.6	1.543	1.659	1.871	2.192	2.638	3.226	3.975	4.907	6.045	7.417	9.052
0.7	2.192	2.398	2.731	3.212	3.865	4.717	5.798	7.143	8.792	10.788	13.181
0.8	2.96	3.282	3.772	4.46	5.382	6.578	8.095	9.987	12.316	15.152	18.574
0.9	3.845	4.314	5.002	5.952	7.216	8.853	10.933	13.537	16.759	20.707	25.508
1.0	4.847	5.498	6.431	7.707	9.399	11.592	14.387	17.904	22.284	27.688	34.309

As another fictitious example, suppose that the model has two continuous parameters P_1 , P_2 , and values of both of them can be between 0 and 1. If the required accuracy for each of the parameters is taken 0.1, then for each one the values 0.0, 0.1, 0.2, 0.3, ..., 0.9, 1.0 can be used, thus $11^2 = 121$ assignments of parameter values have to be evaluated for P_1 and P_2 :

$$\begin{aligned} &\{ (0.0, 0.0), (0.0, 0.1), (0.0, 0.2), \dots, \\ &(0.1, 0.0), (0.1, 0.1), (0.1, 0.2), \dots, \\ &\dots (1.0, 0.8), (1.0, 0.9), (1.0, 1.0) \}. \end{aligned}$$

So, for each of these assignments the error has to be determined (see Table 14.4) and then the set with least error chosen.

The for each of these assignments the error has to be determined (see Table 14.4) and then the set with least error chosen.

The above table illustrates the error for each of these 121 assignments of parameter values. The error is minimal when (P_1, P_2) is assigned values (0.3, 0.1). Thus, the exhaustive search method suggests 0.3 and 0.1 as values for P_1 and P_2 .

A problem with exhaustive search is that for more parameters and smaller grain size it becomes computationally infeasible. For instance, for the simple example described in Sect. 14.3 with 14 parameters it is practically impossible to get a result for grain size 0.1, which even is not a very good accuracy.

Advantages of exhaustive search

- Exhaustive search is very simple to implement.
- It will always find the best set of parameters (global optimum) if all the possible parameter value combinations are explored.

- Exhaustive search is useful as ‘baseline’ method when benchmarking other algorithms.

Disadvantages of exhaustive search

- The main disadvantage of the exhaustive search is that for many real-world problems the set of candidate value assignments is prohibitively large. Its computational cost is proportional to the number of value assignments, which in many practical problems tends to grow very quickly as the size of the problem increases, in particular when the number of parameters is large and/or a high accuracy is needed.

Therefore, exhaustive search is typically used as the best method when the problem size (e.g., the number of parameters and required accuracy) is limited; in other cases this method may easily become too inefficient.

14.5 Parameter Estimation by Gradient Descent

Since exhaustive search is not feasible when the search space of value assignments to parameters becomes too large, other, heuristic methods have been developed for which only part of the search space is explored, where the heuristic used takes into account the error for the given value assignment and the error for alternative choices for this value assignment. The idea of such a heuristic method is that after one point of the search space has been explored and some error value has been found, the neighbourhood of this point is explored to find out in which direction a next point in the search space should be considered. Usually this is the point for which most decrease in error is expected. One of these methods is gradient descent (also called steepest descent or hill climbing). It takes into account how sensitive the error is for the different parameters; in each step the parameters with highest sensitivity are changed most. The sensitivity S_P of the error for a parameter P is based on the difference between the error $\text{error}(P)$ for a value for P and the error $\text{error}(P + \Delta P)$ for a slightly different value $P + \Delta P$ for P . Figure 14.5 shows the basic idea of the sensitivity of a parameter for the error. Here ΔP is a small change of the value of a parameter P .

The sensitivity S_P is defined as

$$S_P = (\text{error}(P + \Delta P) - \text{error}(P)) / \Delta P$$

where the values of the other parameters are kept constant; note that this sensitivity depends on these values of the other parameters.

The idea is that sensitivities can be used to determine from a given value assignment an improved value assignment to the parameters, thus making a specific, informed step in the space of value assignments by which the error is decreased. An analogy of gradient descent is a hiker trying to descend in thick fog

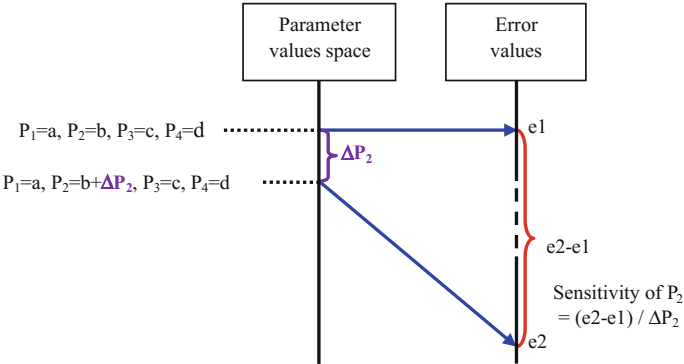


Fig. 14.5 Sensitivity of a parameter

from a mountain to the bottom of a valley by taking at each point the steepest descending path. Gradient descent starts from a random set of parameter values, and at each step, changes the set of parameter values a little, in the direction that decreases the error most. Here each parameter P has to be changed in relation to the sensitivity S_P of the error to the value of P . In this method, all parameters are changed in parallel; parameters with higher sensitivity will change more. In more detail the process goes as follows. By determining the sensitivity of the parameters with respect to the error, the parameters can be identified that are most suitable for adjustment in such a step: if the sensitivity of a given parameter is high, and the error is not small, the current value of that parameter is a good candidate to be changed, and not the value of a parameter of which the sensitivity is low. The aim is to use sensitivities to decrease the value of the error. Therefore, if the sensitivity of a parameter is positive, the value of that parameter needs to be decreased, and if the sensitivity is negative the value of that parameter needs to be increased. Because sensitivities depend on the given value assignment, in the gradient descent parameter estimation method, the sensitivities for each parameter are determined over and over again, in each iteration. When a point is reached in which the error is minimal, all sensitivities will be 0 or very close to 0, as then the differences error $(P + \Delta P) - \text{error}(P)$ will be very small. Therefore during the process the sensitivities will converge to 0; this means that if the sensitivity is a bit high for a given parameter the value for this parameter is not yet optimal, and therefore it should be changed. In summary, the gradient descent method proceeds according to the following steps:

- First an initial assignment AP of parameter values is selected. This can be chosen either totally in random or based on some heuristic knowledge. The time needed to find the most suitable parameter values will also depend on the quality of the selected initial values.
- The model is used to generate a simulation trace for the parameter values from AP

- For the generated trace the error value $\text{error}(\text{AP})$ is determined.
- For any parameter from the parameters set its value is changed for a very small amount ΔP , thus obtaining value $P + \Delta P$
- The model is used to generate a simulation trace for the parameter values from AP with P adjusted to $P + \Delta P$
- The new error value is recalculated; say $\text{error}(\text{AP}, P, P + \Delta P)$
- From the difference between the two obtained error values $\text{error}(\text{AP}, P, P + \Delta P)$ and $\text{error}(\text{AP})$ the sensitivity S_P of parameter P in that situation (and relative to the values of the other parameters) is determined as:

$$S_P = (\text{error}(\text{AP}, P, P + \Delta P) - \text{error}(\text{AP})) / \Delta P.$$

In this way the sensitivity of each parameter with respect to the error is determined separately. If a small change in parameter value leads to a large difference in error, that means the sensitivity of that parameter for the error is very high.

- A new assignment AP of values to the parameters is obtained by changing the value of each parameter P in proportion to its sensitivity, with η as rate of change:

$$\text{new } P = P - \eta * S_P.$$

Here η is the adjustment rate, which defines the size of the changes in each step. If η is too small, convergence is needlessly slow, whereas if η is too large, the adjustment process will overshoot and can even diverge. Here, it is supposed that the adjustment rate η is a fixed constant. However, it can also change in each step, as shown in Chong and Zak (2013).

```

Choose initial assignment of parameter values AP %Arbitrary starting point
Set adjustment rate  $\eta$  %Adjustment rate choice
Set a small  $\Delta$  for parameter value difference
Run the model for AP to generate a trace %Generate simulation trace for AP
Calculate  $\text{error}(\text{AP})$  %Determine error for AP
While (any significant change has occurred in AP in the last step)
  For each of the parameters  $P$  in AP %Determine the sensitivity of the error
    Run the model for  $\text{AP}(P, \Delta)$  where the value  $V$  of  $P$  in AP is replaced by  $V + \Delta$ 
    Calculate  $\text{error}(\text{AP}(P, \Delta))$  %Determine error for  $\text{AP}(P, \Delta)$ 
    Calculate sensitivity  $S_P = (\text{error}(\text{AP}(P, \Delta)) - \text{error}(\text{AP})) / \Delta$ 
  For each of the parameters  $P$  in AP
    Set new value  $P = P - \eta * S_P$  %Adjust AP
  End for
End while

```

Box 14.2 Pseudo-code for gradient descent

The gradient descent method has been implemented for the example described in Sect. 14.3 above and performed for different initial values of the parameters. In Fig. 14.6 the solution found is shown for one case (upper graph) and the lower graph shows the error for the different iterations for this case.

Advantages of gradient descent

- Gradient descent is relatively simple.
- It is a fast method; in a few iterations, it can find a local minimum.
- With certain assumptions, convergence to a local minimum can be guaranteed.

Disadvantages of gradient descent

- The outcome of gradient descent may depend on the initial point
- This method can get stuck in local minima
- Many calculations have to be done in each iteration

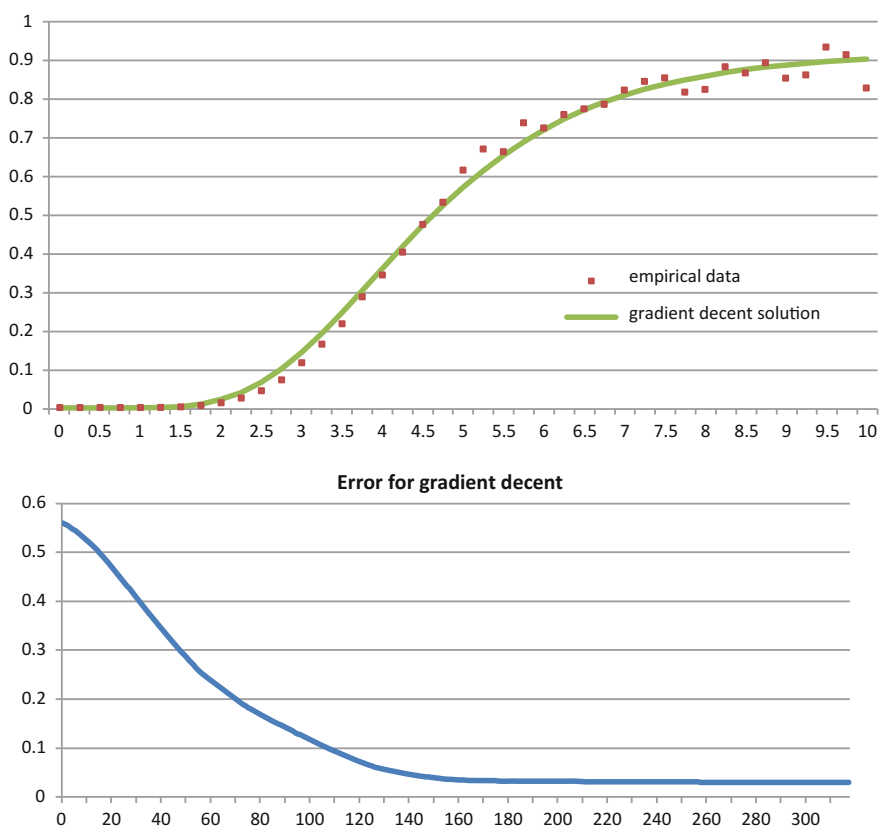


Fig. 14.6 A solution found by gradient descent for the example, and the error at each point during gradient descent against number of iterations

- There are situations in which gradient descent can be slow and inefficient. To overcome such problems, a number of variations on gradient descent have been developed, such as conjugate gradient descent (Chong and Zak 2013).

14.6 Parameter Estimation by Random Gradient Descent

In gradient descent (or hill climbing), in each step the parameter values are changed a little, in a way that this small change causes the largest possible decrease in error. Therefore, the sensitivity of the error to each of the parameters has to be computed again and again, which in some cases is a time consuming process. There are a number of other descent methods that do not necessarily follow the line of steepest descent. Random gradient descent (or stochastic hill climbing) is one of these methods, which can be useful when the gradient is too time consuming to compute all the time. In random gradient descent, a slight random change of the current assignment of parameters values is applied and the result is accepted only if the new value of the error lies below the current value. Otherwise, a new change to the parameter values is tried (Baldi and Brunak 2001). In fact, in this method any change that causes a decrease in error is accepted (not necessary the largest one). In the hiker metaphor, in this case the hiker will choose any path that goes downward, not necessarily the steepest one. So, in each step the gradient descent method moves in the best way (steepest descent, gradient), but, the random gradient descent method just moves in a good way (not necessary the best one).

```

Set  $\alpha$  = neighbourhood limitation           % A value like 0.1
Choose any initial assignment of parameter values AP   %Arbitrary starting point
Run the model for AP to generate a trace                %Generate simulation trace for AP
Set E = error(AP) for this trace                       %Determine_error for AP
While E > maxerror
    Generate AP' from AP by for each parameter value P in AP
        Set  $P' = P + \alpha * (\text{rand}() - 0.5)$            % Choose AP' in the neighbourhood of AP
    Run the model for AP' to generate a trace            %Generate simulation trace for AP'
    Set E' = error(AP') for this trace                  %Determine_error for AP'
    If (E' < E )                                         %If error for AP' is smaller
        Set AP = AP'                                   %Choose AP' as new AP
        Set E = E'
    End if
End while
    
```

Box 14.3 Pseudocode for random gradient descent

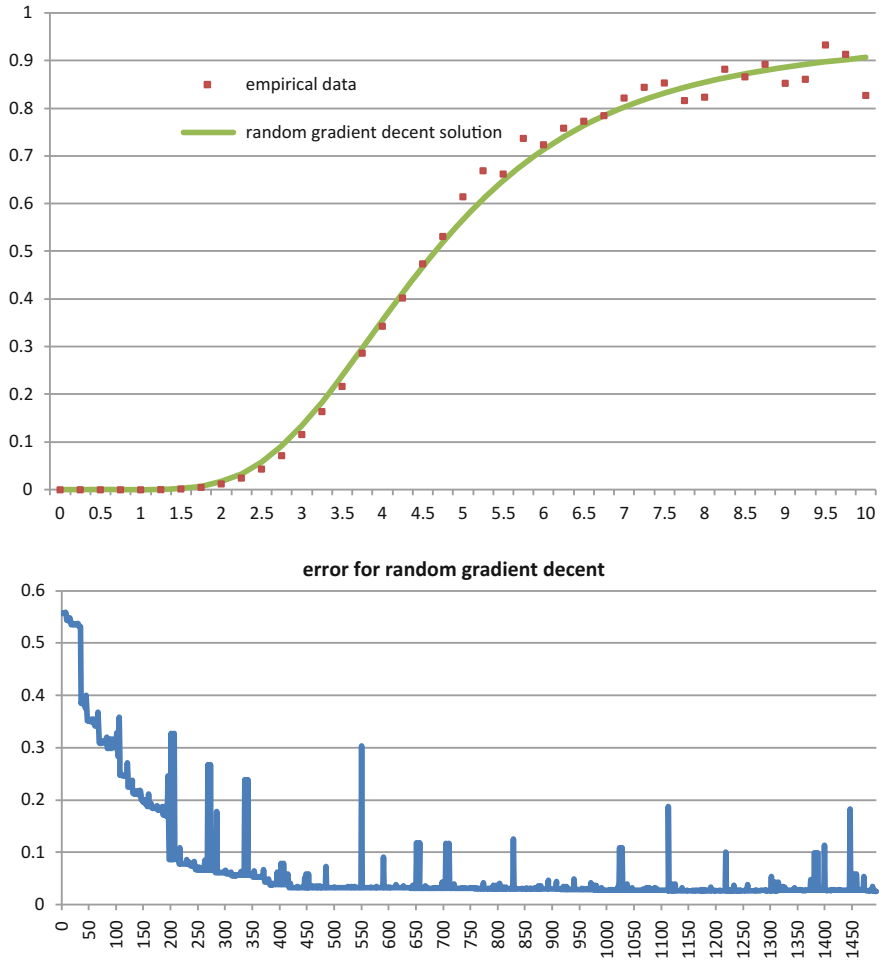


Fig. 14.7 The error at each point during random gradient descent

The random gradient descent method has been implemented for the example which is described in Sect. 14.3. Figure 14.7 shows the error for the different iterations. Compared to what is shown in Fig. 14.6 for gradient descent the number of iterations to reach a low error level can be 10 times more.

Advantages of random gradient descent

- Random gradient descent is relatively simple
- Not many calculations have to be done in each iteration; no sensitivities have to be determined
- With certain assumptions, convergence to a local minimum can be guaranteed.

Disadvantages of random gradient descent

- Many iterations may be needed before it reaches a small error
- The outcome of gradient descent may depend on the initial point
- This method can get stuck in local minima

14.7 Parameter Estimation by Simulated Annealing

A gradient descent method never makes moves toward states with a higher error value; therefore it is guaranteed to at least find a locally optimal parameter value assignment, with a local minimum of the error function. In contrast, in a purely random walk, the parameter values are randomly changed. Like exhaustive search such a random method is complete, in the sense that it will always find a global minimum in the end, but is extremely inefficient. Therefore, it may be reasonable to try to combine gradient descent with a random walk in some way in order to achieve both acceptable efficiency and completeness. Simulated Annealing is such a method. Actually, the most important advantage of this method (in comparison to gradient descent) is that the simulated annealing method can pull the search process out of local minima.

Simulated annealing is inspired by physical annealing in metallurgy, where physical substances are heated and melted, and then gradually cooled down until some solid state is reached. In this process, the goal is reaching a state in which the substance has a minimum of energy. In metallurgy, this goal will be attained if the substance is cooled down in a sufficiently slow manner.

The notion of slow cooling down in physical annealing is implemented in the simulated annealing method as a slow decrease in the probability of accepting worse solutions while it explores the search space. This property is controlled by a variable called Temperature (T). Accepting a worse assignment of parameter values (which leads to higher value of error) is a fundamental property of simulated annealing because it allows for a more extensive search for the optimal solution.

Simulated annealing starts from a random assignment of parameter values AP . At each step, the method considers a neighbouring assignment of parameter values AP' of the current set AP , and probabilistically decides between moving to set AP' or staying in AP . These probabilities ultimately lead the system to move to better sets of parameter values (with lower error).

Inspired from physical annealing, in this method there is a variable called temperature (indicated by T), which has a value decreasing (cooling) during the process. Actually, the parameter temperature controls the probability of doing 'downhill' actions during the process. When the value of this parameter is high (at

the beginning) the probability of accepting a new set of parameter which leads to higher error is high. During the progress of the method, this value will decrease, and when the value of this variable is very low (at the end), the probability of accepting such a new set is almost zero. Because of that, simulated annealing can pull out of local minima and find the globally most optimal point. If the temperature is cooled down in a sufficiently slow manner, it will find the global minimum (the best possible assignment of parameter values).

It starts from a random set of parameter values AP^0 and generates a succession of parameter values $AP^1, AP^2, \dots$ in order to decrease the error. New candidate sets are generated around the current set of parameter values by slightly changing these values in a random way for each parameter. For each iteration i the new values are uniformly distributed in intervals centered around AP^i . If the new assignment of parameter values has a decreased the error level $\text{error}(AP^{i+1})$, in comparison to the error level $\text{error}(AP^i)$ of the previous assignment, it will be accepted (i.e., when $\Delta E = \text{error}(AP^{i+1}) - \text{error}(AP^i) \leq 0$). Otherwise, the new set will be accepted with probability $e^{-\Delta E/T}$, which depends on the temperature and the difference between previous error and the new one.

```

If  $\Delta E \leq 0$ 
    accept the new point: take  $AP^{i+1}$ 
else
    accept the new point with probability  $e^{-\Delta E/T}$ 
end if

```

Here T is the temperature and $\Delta E = \text{error}(AP^{i+1}) - \text{error}(AP^i)$.

So, a new assignment of parameter values which generates a higher level of error, will be accepted with probability of $e^{-\Delta E/T}$. This probability is only dependent on ΔE and T . If ΔE is very small (the new assignment of parameter values increases the error level only a bit), or if the temperature is very high, then the new assignment of parameter values will be accepted with a high probability. This probability will decrease by decreasing the temperature or increasing ΔE .

Figure 14.8 shows $e^{-\Delta E/T}$ as a function of ΔE when $T = 1$. Figure 14.9 shows $e^{-\Delta E/T}$ as a function of T when $\Delta E = 2$.

As can be seen, the value of the function $e^{-\Delta E/T}$ is close to 1 when ΔE is very small or T is large. On the other hand, it is less when ΔE is large and T is small.

The simulated annealing method starts at a user defined temperature T_0 and the temperature will be decreased in each iteration. The process is terminated when the temperature is so low that no more significant improvement can be reached.

Actually, the core of the simulated-annealing method is quite similar to the gradient descent method. But, instead of picking the best move it picks a *random* move. If the move improves the situation (decreases the error), it is always accepted (like the random gradient descent). Otherwise, the algorithm accepts the move to a

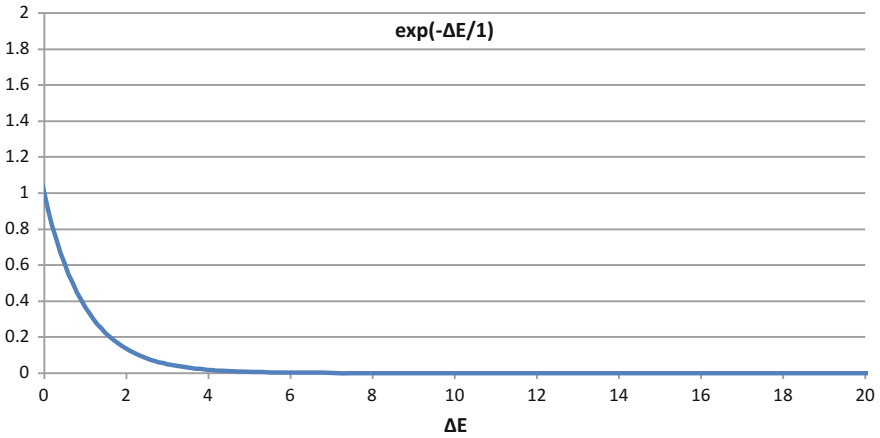


Fig. 14.8 The values of $\exp(-\Delta E./T)$ depending on ΔE for $T = 1$

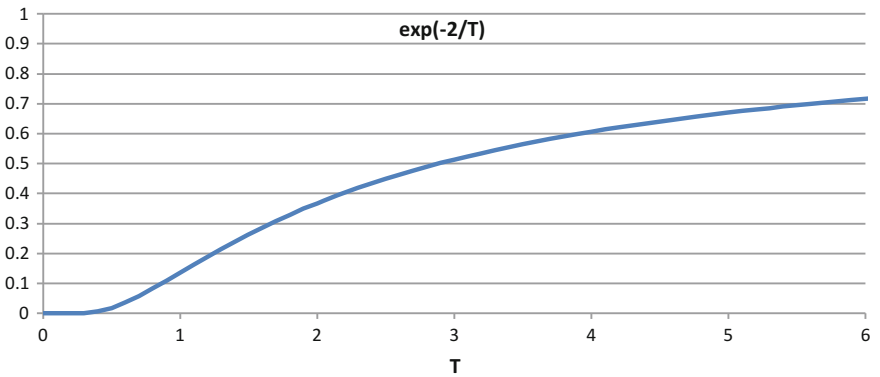


Fig. 14.9 The values of $\exp(-\Delta E./T)$ depending on T for $\Delta E = 2$

higher error with some probability less than 1. The probability decreases exponentially with how bad the move is: the amount ΔE by which the error is increased. The probability also decreases as the temperature T goes down: bad moves are more likely to be allowed at the start when T is high, and they become more unlikely as T decreases. If the temperature T decreases slowly enough, the method will find the best set of parameters with probability approximating 1.

```

Set  $\lambda$  = cooling rate of temperature           % a value between 0.9 and 1.0, like 0.995
Set  $\alpha$  = neighbourhood limitation           % a value like 0.1
Set initial T                                % Initialization
Set initial AP randomly                      % Initialization
Run the model for AP to generate a trace      % Generate simulation trace for AP
Set E= error (AP) for this trace              % Calculate the error for assignment AP
While (T  $\geq$  Tmin)
    Generate AP' from AP by for each parameter value P in AP
        set P' = P +  $\alpha$  * (rand() - 0.5)    % Choose a neighbouring assignment AP'
    Run the model for AP' to obtain a trace    % Generate simulation trace for AP'
    Set E' = error(AP') for this trace         % Calculate the error for assignment AP'
     $\Delta E = E' - E$ 
    If ( $\Delta E < 0$ )                           % If the new assignment AP' reduces the error
        AP = AP'                             % change assignment to one with lower error
        E = E'                               % update the error value
    else                                       % if the assignment AP' does not reduce the error
        if random() > exp(- $\Delta E / T$ )         % do the next step with the probability exp(- $\Delta E / T$ )
            AP = AP'                         % change assignment to one with higher error
            E = E'                           % Update the error value
        end if
    end if
    T = T *  $\lambda$                              % Decrease the temperature
end while

random() = a function that generates a random value between 0 and 1

```

Box 14.4 Pseudo-code for simulated annealing

As mentioned, the temperature T is decreased during the progress; the cooling schedule defined by parameter λ is of paramount importance for the performance of simulated annealing. There is a trade-off between the quality of the final solution and the execution time, the latter being sensitive to the speed of the temperature decrease. Here, it is done by multiplying T in each iteration with λ which is a number < 1 (and usually > 0.95). If a very high initial temperature T is chosen, there will be a waste of computational resources. In the contrary case of low initial T , the process could get caught in assignments of parameter values which are not the best ones. It is very hard to establish a general rule for determining the ideal initial temperature.

The simulated annealing method has been implemented for the example described in Sect. 14.3. Figure 14.10 shows the error for the different iterations.

Advantages of simulated annealing

- Simulated annealing is effective in finding a good assignment of parameter values for models with a huge numbers of parameters.
- Although the final point is not deterministically guaranteed to be the best one, if the temperature decreases in an appropriate pace, it will find a very good answer.

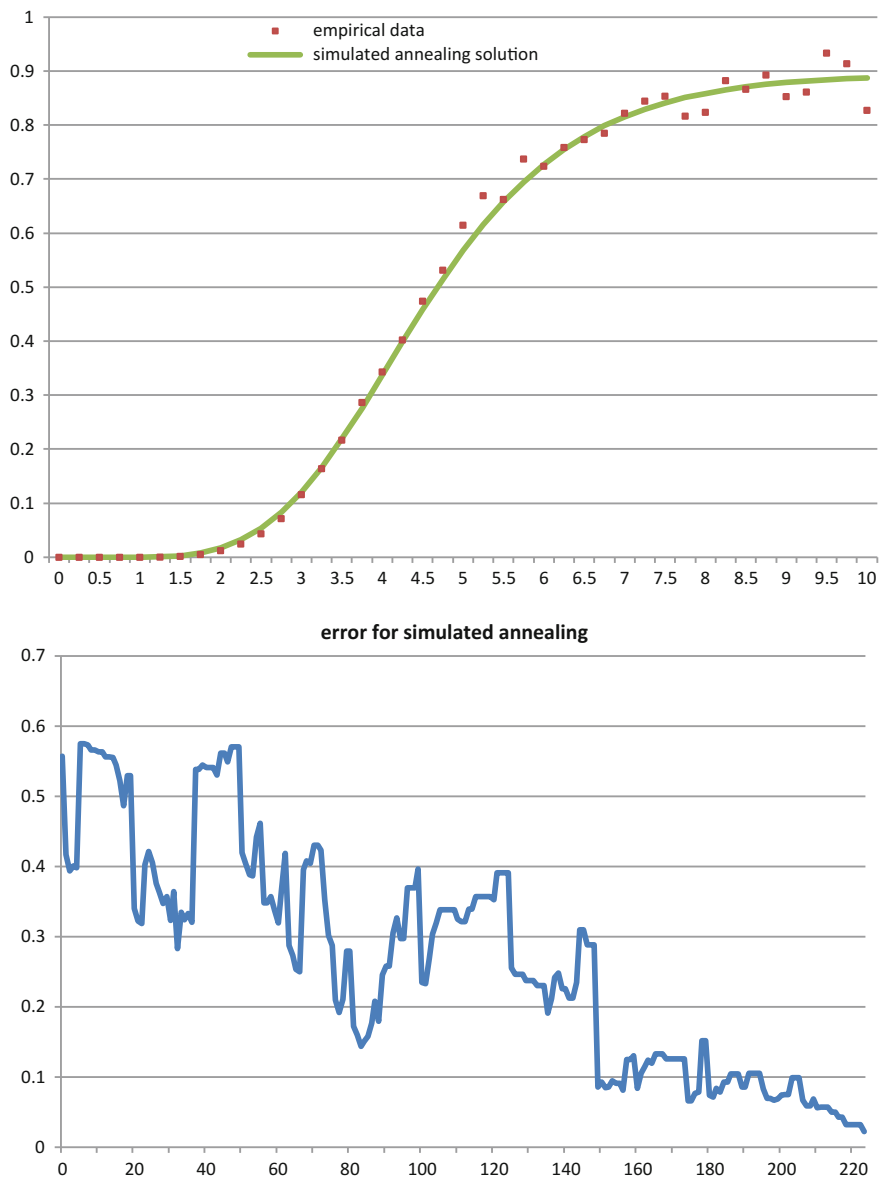


Fig. 14.10 The error for the different iterations in simulated annealing

Disadvantages of simulated annealing

- In comparison to gradient descent or random gradient descent, it is slow. This method needs more iterations to converge to an optimal answer.

14.8 Discussion

This chapter is based on work with Amin Tabatabaei and Dilhan Thilakaratne. For example, all data and figures used were generated by them. In this chapter only some of the most often used parameter estimation methods have been discussed. Sometimes variants of these methods or other methods are used; for example, see van den Bos (2007), Chong and Zak (2013), Thilakaratne (2015). As can be expected the different parameter estimation methods and settings found different solutions for the example model. In Table 14.5 the parameter values are shown as found by the processes displayed in Figs. 14.5, 14.6, and 14.9; recall that all of these solutions had error 0.03 or just below that value.

This illustrates that usually there is not one unique best solution (related to a global optimum) that is the only relevant solution, but multiple solutions (related to local optima) are possible and relevant. It may depend on the context which of these solutions are most relevant. In practice, to simply go for some unique solution related to a global optimum, as sometimes may be suggested, will often not be recommendable. Instead, better insight will be obtained when an overview is found of different solutions relating to local optima with their respective errors, especially when these errors are not that different, as in the above case.

In recent work an approach is introduced in which an error function is defined for any requirement expressed as a dynamic property as discussed in Chap. 13. This error function is based on approximate satisfaction, which provides a measure for how close to satisfaction a requirement is. It generalises the error function based on least squares of residuals when simulated values are compared to empirical values as described in Sect. 14.2.4 to the case of arbitrary requirements.

Table 14.5 Different solutions found for error 0.03

	Gradient decent	Random gradient decent	Simulated annealing
ω_{es_b,ws_b}	0.64	0.92	0.59
ω_{ws_s,ss_s}	0.82	1.00	0.52
ω_{ws_b,ss_b}	0.71	0.90	0.97
ω_{ss_s,srs_s}	0.85	1.00	0.55
σ_{srs_b}	2.84	2.79	6.94
τ_{srs_b}	0.01	0.10	0.01
ω_{ss_b,srs_b}	0.69	0.78	0.73
ω_{ps_b,srs_b}	1.00	1.00	1.00
σ_{ps_b}	6.58	6.59	5.00
τ_{ps_b}	0.10	0.24	0.10
ω_{fs_b,ps_b}	0.89	1.00	0.97
ω_{srs_s,ps_b}	0.88	1.00	1.00
ω_{srs_b,fs_b}	1.00	1.00	0.90
ω_{ps_b,es_b}	0.66	0.62	0.50

References

- P. Baldi, S. Brunak, *Bioinformatics: The Machine Learning Approach* (MIT Press, 2001)
- E.K.P. Chong, S.H. Zak, *An Introduction to Optimization* (Wiley, 2013)
- A. Damasio, *The Feeling of What Happens: Body, Emotion and the Making of Consciousness* (Harcourt Brace, 1999)
- A. Damasio, T.J. Grabowski, A. Bechara, H. Damasio, L.L.B. Ponto, J. Parvizi et al., Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nat. Neurosci.* **3**, 1049–1056 (2000)
- E. Gilbert, K. Karahalios, Predicting tie strength with social media, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI'09*, pp. 211–220 (2009)
- M. Granovetter, The strength of weak ties: a network theory revisited. *Sociol. Theory* **1**, 201–233 (1983)
- M.J. Hove, J.L. Risen, It's all in the timing: interpersonal synchrony increases affiliation. *Soc. Cognit.* **27**, 949–960 (2009) (doi:[10.1521/soco.2009.27.6.949](https://doi.org/10.1521/soco.2009.27.6.949))
- P.V. Marsden, K.E. Campbell, Measuring tie strength. *Soc. Forces* **63**, 482–501 (1990)
- C. Moler, Least squares, in *Numerical Computing with MATLAB, Chapter 4. Society for Industrial and Applied Mathematics* (2004)
- J. Parvizi, G.W. Hoesen, J. van Buckwalter, A. Damasio, Neural connections of the posteromedial cortex in the macaque: implications for the understanding of the neural basis of consciousness. *Proc. Natl. Acad. Sci.* **103**(5), 1563–1568 (2006)
- E. Pearce, J. Launay, R.I.M. Dunbar, The ice-breaker effect: singing together mediates fast social bonding. *Roy. Soc. Open Sci.* (2015). doi:[10.1098/rsos.150221](https://doi.org/10.1098/rsos.150221)
- V. Strejc, Least squares parameter estimation. *Automatica* **16**(5), 535–550 (1980)
- D.J. Thilakarathne, A parameter estimation method for dynamic computational cognitive models, in *Proceedings of the 6th Annual International Conference on Biologically Inspired Cognitive Architectures, BICA'15, Procedia Computer Science*, vol. 71 (2015) pp. 133–142
- A. van den Bos, *Parameter Estimation for Scientists and Engineers* (Wiley Interscience, Hoboken, 2007), p. 2007

Part V
Philosophical, Societal and Educational
Perspectives

Chapter 15

We Don't Believe in Ghosts, Do We?

What Is It that Drives Dynamics

Abstract In this chapter it is discussed how dynamics has been a challenging issue in different disciplines since long ago. This issue has been addressed for different domains, in Physics but also in Mathematics, Cognitive Science and Philosophy of Mind. In the development of Physics it has led to notions such as velocity, momentum, kinetic energy and force that drive motion in mechanics. The issue of dynamics is still out there today, for example, in the domain of Cognitive Science and Philosophy of Mind concerning the physical realism of assumed but not directly physically observable mental states such as desires and intentions that are supposed to drive (physically observable) behaviour. Four cases of dynamics within different traditional disciplines are discussed in this chapter. Similarly, it is shown how causal graphs and transition systems (often used in AI and Computer Science) can be interpreted from a perspective of dynamics. The chapter provides a unified view on the explanation of dynamics across different disciplines. This view is related to the basic assumptions underlying the Network-Oriented Modeling approach based on temporal-causal networks.

15.1 Introduction

The Network-Oriented Modeling approach based on temporal-causal networks discussed in this book involves an important continuous-time temporal dimension. This temporal dimension makes it a dynamic modeling approach; as discussed in Chap. 2, Sect. 2.9 it covers all smooth continuous dynamical systems. Dynamics has puzzled researchers since long ago. Among them are Greek philosophers such as Zeno of Elea (about 490–425 BC) and Aristotle (384–322 BC). They pointed at the phenomenon that the world occurs to us in different states at different points in time. However, for the transition from a given physical state to another physical state, it is not always clear from the given physical state what will be different in the next state. For example, Zeno and Aristotle argue that at one specific instant in the physical world (a snapshot) a moving arrow cannot be distinguished from an arrow in rest, yet the next state for a moving arrow is different (e.g., Heath 1931).

What is it in this given state that is driving the change to a next state in one case but which apparently is absent in the other case? When no physical property can be found in the given original physical state that can explain this change, what other entity can be there to explain the change? Usually an entity that is not part of physical reality, and therefore cannot be sensed in any way, but still may bring about changes in the physical world, is called a ghost. If for a transition from a given physical state nothing physical can be found in this state that can explain what will be different in the next state, then it may seem that this change has to be attributed to a ghost or ghost-like entity or property in the original state.

So, to explain dynamics that clearly occurs in the world, do we have to believe in ghosts or ghost-like state properties? This issue has been discussed since the time of Zeno and Aristotle for different domains. In fact the issue is still out there today, for example, in the domain of Cognitive Science and Philosophy of Mind (e.g., Kim 1996) concerning the physical realism of assumed but not directly physically observable mental states such as desires and intentions, that are supposed to drive (physically observable) behaviour. But also in the development of Physics this issue has played an important role and has led to notions such as velocity, momentum, kinetic energy and force that drive motion in mechanics, notions which just as a property like 'being moving' of an arrow cannot be observed in a snapshot at one specific instant; e.g., René Descartes (1596–1650), Christiaan Huygens (1629–1695), Isaac Newton (1643–1727) and Gottfried Wilhelm Leibniz (1646–1716). Similarly, in Mathematics abstract notions such as derivatives, differential equations and Taylor series for mathematical functions have been introduced addressing the question what drives dynamics.

To address the issue discussed above, during history often solutions have been proposed by assuming hypothetical state properties called *potentialities* or *anticipatory state properties*. These are state properties p that, when they occur in a state, indicate that a specific property a will be different in the next state. As they are hypothetical, a next question then is if and how they can be related to other, more genuine state properties. Assuming potentialities, it is possible to explain from a given state why and how a next state is different from this given state. But potentialities themselves can still be ghost-like properties that are not part of the physical world. Being ghost-like is not a very satisfactory status, and certainly it is not when the occurrence of such a ghost is unpredictable.

There are two ways out of this. The first option is to find a physical state property c (a *realiser*) that within a state always co-occurs with a given potentiality p , so that the potentiality is at least equivalent to a physical state property. This is one way in which in some cases a certain extent of realism (and through this predictability) can be attributed to a potentiality. However, this is not always an option; for example, in the case of a moving arrow such a physical state property c simply does not exist. Then the other option is that a potentiality is not equivalent to any physical state property in the state in which it occurs, but it still has (temporal) relations to physical state properties in previous states. Such temporal relations may also provide a possibility to predict the occurrence of a potentiality. So, whereas in the

given state the potentiality might still be considered as ghost-like, at least the occurrence of this ghost is predictable from previous states.

The second option is usually applied within Physics for notions such as velocity, momentum and kinetic energy. For the notion of force, in general it is possible to apply the first option, although a drawback is that in this case the realising state property c is not unique but depends on context: a given force f can be realised in different ways. A realiser c for f may be related to a context of either electricity, magnetism, elasticity, gravitation, or ..., or even to a combination of such contexts.

Within Philosophy of Mind the first option is applied in the context of reduction relations for mental states based on physical realisers (e.g., Kim 1996, 1998); also here the issue of context dependency may occur, as different organisms may use different realisations for similar mental states. The second option is also considered within Philosophy of Mind to be relevant for a certain type of mental states (see Kim 1996, pp. 193–207); the second option is sometimes described in that context by a (temporal) relational specification of a mental state (e.g., see Kim 1996, pp. 200–202). Such a temporal relational specification is in fact exactly how in Physics the second option is used: by a temporal relational specification of state properties such as velocity, momentum and kinetic energy.

In recent years, within Cognitive Science, dynamics has been recognized and emphasized as a central issue in describing cognitive processes; for example, Port and Gelder (1995); see also Chap. 1. As has been discussed in more detail there, the notion that lies at the heart of this dynamical perspective is the notion of a state-determined system. This type of system is based on the assumption that (properties of) a given state fully determine (the properties of) future states. This explicitly puts the focus on the properties of the previous state to explain why some change has occurred in a next state. Therefore state-determined systems can be explained well in terms of potentialities.

Four cases within different disciplines (Cognitive Science, Physics, Mathematics, Computer Science) are analysed in this chapter and show how in history the notion of potentiality has led to a number of often used concepts within these disciplines. Among them are concepts like desire and intention in Cognitive Science, momentum, kinetic energy, and force in Physics, and derivatives of a function and Taylor approximations in Mathematics. Similarly, causal graphs and transition systems (often used in Computer Science) can be interpreted from the perspective of potentialities. This unified view on the explanation of dynamics across different disciplines is also one of the contributions of the chapter.

The chapter is structured as follows. Section 15.2 discusses the perspectives of Zeno and Aristotle, thereby illustrating the problem of explaining changed states for nonliving entities in more detail. Section 15.3 does the same for living entities such as animals, and humans in particular. In Sect. 15.4 the notion of potentiality and its actualisation is discussed in some more detail. In Sect. 15.5 this fundamental notion is addressed for a case study on potentialities for (loco)motion in Physics. More specifically, it is discussed how in classical mechanics as developed by Descartes, Huygens, Newton and Leibniz, among others, a potentiality for ‘*quantity of motion*’ and one for ‘*moving force*’ were developed; in modern physics known as

momentum and kinetic energy. Both can be related to the concept of velocity. In Sect. 15.6 it is shown that the mathematical formalisation of this concept velocity (based on the mathematical notion derivative) is inherently based on properties of states at different points in time. Section 15.7 describes how potentialities are involved in models described by causal relations and transition systems, which is a well known representation within Artificial Intelligence and Computer Science to specify change in systems.

An issue considered within Cognitive Science and Philosophy of Mind is whether mental state properties are genuine state properties. For example, as a desideratum it might be posed that they should be identifiable with 'real' and perhaps even directly observable state properties. This issue of realism also applies to potentialities; this is discussed in Sect. 15.8. Assuming that potentialities exist as anticipatory state properties that can explain properties of subsequent, changed states, a next question is how the occurrence or change of a potentiality itself can be explained. In Sect. 15.9 it is discussed how a changed potentiality can be explained by a higher order potentiality (i.e., a potentiality to get a potentiality). In Sect. 15.10 it is discussed how interaction between objects can be interpreted as a transfer of potentialities. It is argued that, if higher-order potentialities are generally assumed to explain changed potentialities, changes due to interaction between objects have to be attributed to higher-order potentialities as well, i.e., such an interaction is characterised by the higher-order potentialities that co-occur with it. Historically, the second-order potentiality for motion was formalised by the notion of 'force' in classical mechanics; interactions between objects are characterised by the forces they impress on each other. Section 15.11 addresses the issue of realism in more detail, in particular for cases in which multiple relations to reality are involved. Section 15.12 discusses how the notion of potentiality is a basic notion underlying state-determined systems. Finally, Sect. 15.13 is a discussion.

15.2 Is Motion of Nonliving Entities Driven by Ghosts?

Following Zeno of Elea (about 490–425 BC), this section first discusses in Sect. 15.2.1 why changed states cannot always be explained on the basis of given 'real' state properties, and, hence, thus providing an argument for why dynamics does not exist in reality. Next, in Sect. 15.2.2 to solve this problem the option to add anticipatory state properties called potentialities to obtain an explanation of changed states are discussed.

15.2.1 *Zeno About Arrows that Are Moving and Unmoving*

An arrow moving from A to B traverses a number of positions between A and B at different time points. Zeno asked himself what the difference in state is between a

moving arrow in a certain position P at some time point t and an arrow at rest in t at the same position P . He came to the answer that there is not any state property that differs for the two states, so there is no difference in t between a moving arrow and an arrow at rest. From this he concluded that in reality motion does not exist: it is just an illusion, made up by human perception and processing. Zeno summarised his view in the following paradoxal formulation:

If everything is either at rest or moving when it occupies a space equal to itself, while the object moved is in the instant, the moving arrow is unmoved.

This formulation was taken from Aristotle (translated by Heath 1931) who incorporated some of Zeno's work in his writings; a book written by Zeno himself unfortunately disappeared (it is said to be stolen from Zeno). Also Aristotle (384–322 BC) claimed that motion and change do not refer to anything existing in reality:

Again, there is no such thing as motion over and above the things. It is always with respect to substance or to quantity or to quality or to place that what changes changes. But it is impossible, as we assert, to find anything common to these which is neither 'this' nor quantum nor quale nor any of the other predicates. Hence neither will motion and change have reference to something over and above the things mentioned, for there is nothing over and above them. [from Aristotle, 350 BC a, Physics, Book III, Part 1]

In the 19th and 20th century a number of technological developments made it possible to do further experiments. For example, the concept of movie was developed and the technical equipment to implement this concept; a movie is created just by successively displaying a large number of static pictures (e.g., 24/s); see, for example, Burns (2000). Nevertheless humans watching a movie get the impression of motion just like in reality, from which the word 'movie' stems. But nothing really moves in a movie, which is a paradox in words in the same spirit as Zeno's 'the moving arrow is unmoved'. Other examples of 20th century technological developments that support Zeno's view are television, computer animation and virtual reality. All of these experiences support Zeno's view in the sense that having an impression of motion is not an indication that in reality anything like motion exists at all (Fig. 15.1).

These observations suggest the following view on reality. Instead of motion or change, which are human illusions that do not exist in reality, what does exist is the concept of reality as different *states* that are related by a *succession* relation (and can be labeled by time points). Each of these states can be described by the state properties it has; properties make use of language elements for basic state concepts that together form a *state ontology*.

So, for example, to create a virtual reality, i.e., something artificial similar to reality, it is sufficient to display a succession of states at a frequency of, e.g., 20 pictures per second; nothing at all needs to be moved. Indeed, the more recent development of movies has proven that this works: nothing actually moves in a movie, but we perceive movement. However, for a range of lower frequencies stroboscopic effects occur, that for viewers lead to disorientation, and for a still lower frequency range the pictures are perceived one-by-one as separate pictures.

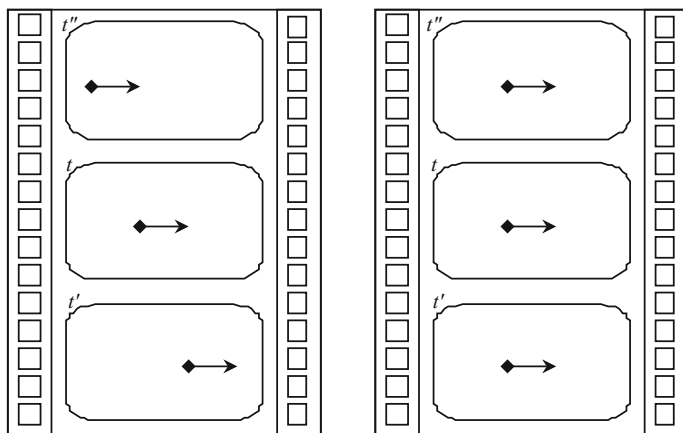


Fig. 15.1 Zeno's two arrows at three different points in time $t'' < t < t'$: the moving arrow (*left hand side*) and the arrow at rest (*right hand side*); at time point t there is no difference in state

If motion or change itself was considered part of reality, properties of successive states could be explained by referring to this change. For example:

Why is the arrow at t' at position P' ?

The arrow is at position P' at t' because

at t it was at position P , and

it was moving in the same direction P' has from P , and

nothing was in its way.

This is almost tautological:

A changed position P' occurs because the position P is changed

If motion as such does not exist in the state ontology to conceptualise reality, such an explanation is unsatisfactory: this explanation makes use of a concept (for 'motion') that is not 'real', i.e., it does not belong to the ontology that conceptualises physical reality. Still the different properties of the states in succession at different points in time exist in reality; they *are* based on the ontology used to conceptualise reality and they ask for an explanation that uses such real properties in preceding states.

More specifically, how to explain for a given state, that for one case, for example with an arrow at rest, a next state has the same property (same position of the arrow), whereas in another case, for example a snapshot of a moving arrow, a next state has a different property (the arrow at a different position)? How can a new property occurring in a next state be explained without involving an unreal concept such as motion? Only taking the usual state properties into account will not suffice to explain differences in properties of subsequent states, because according to

Zeno's analysis the state of an object at rest at time t has exactly the same physical state properties as the state obtained as a snapshot of a moving object at time t . No distinction can be made on the basis of these state properties, and therefore it is inexplicable by physical properties why in one case a subsequent state has a property different from the properties of such a subsequent state in the other case. This seems to contradict the assumption of determinism: given the state at t with its physical properties, the state at t' is not determined, as the position of the arrow can either be still at P or it can be at another position. So, in particular this process cannot be described as a state-determined system, when only the available state properties are used. In how far would it be possible to add state properties beyond these existing ones in order to obtain a state-determined system? This question is addressed in the subsequent sections.

15.2.2 *Adding Anticipatory State Properties to Describe a State: Potentialities*

In order to achieve that the assumption of determinism is satisfied, an additional state property p is needed for the state at t , that makes a difference: if this state property p occurs at t , then the position of the arrow will be different at t' , and if this state property p does not occur it still will be in P at t' . So, assume that to conceptualise a state, in the state ontology more ontological elements can be included for state properties than only the apparent physical ontology and state properties. In particular, a hypothetical or imaginary additional type of *anticipatory state property* or *potentiality* can be assumed in the state ontology: the potentiality p to subsequently get a different state property at a next instant, i.e., the potentiality for the state to become changed. This is a step that may seem rather artificial, but it has been made in the history of science, not only once, but many times. Then the following explanation can be made:

Why is the arrow at t' at position P' ?

The arrow is at position P' at t' because

at t it was at position P , and

at t it had the potentiality p to be at P' , and

at t nothing in the world excluded it from being at P'

So, what is the situation? The arrow can be in two states at time t , which in the physical world cannot be distinguished; these states are exactly identical. But there still is a difference in the state at t that will become observable only at instant t' after t : either at t' the arrow is still in the same position or it is in another position. The difference in state t is whether or not an additional potentiality indicated by state property p occurs at t : in the former case at t no state property p occurs, and in the latter case a state property p does occur at t . This potentiality p is not a normal physical state

property; it cannot be physically observed in the state. But it still makes that the arrow will have a different position at time t' . Typically, such a nonphysical entity showing its presence only by moving an object is called a *ghost*. Therefore, assuming determinism and state-determined systems, the conclusion can be drawn that state-determined dynamics or motion in the physical world is *driven by ghosts*.

During the development of the scientific discipline of Physics in history a number of such ghost-like state properties have been added to the physical ontology, so that nowadays they are treated as more or less normal physical entities, and, for example, people can now say that they believe in science, not in ghosts. This development within Physics is discussed in more detail in Sect. 15.5. Maybe needless to say is that in general the idea that nonphysical ghost-like entities affect the physical world is not appreciated much, as in the physical domain it is preferred that physical effects always have physical causes, which seems to be contradicted by ghost-like state properties causing physical state properties. Moreover, by what kind of mechanism could nonphysical entities affect the physical world? Furthermore, from a practical angle, if a next state becomes predictable from the given state only through such a ghost-like state property, in order for human beings to actually determine such a prediction, it has to be known exactly what ghost-like property is present. So, for practical purposes, if one has to deal with ghost-like state properties an important question becomes: how can these ghost-like properties themselves be predicted? Having to live with such ghost-like entities is one thing, but let them at least be predictable then! Within Physics such predictability has been achieved. This problem of *ghost-like causation* will be addressed in subsequent sections. But first in Sect. 15.3 the idea of motion in living entities is discussed.

15.3 Is Motion of Living Entities Driven by Ghosts?

For motion of living entities such as animals an analysis has been made that has some similarity to the analysis for nonliving entities discussed in Sect. 15.2.1.

15.3.1 *Mental States Driving Motion*

Often used explanations of animal (or human) actions refer to internal mental states. For example, for a creature B which has the capability to move:

Why is B at t' at position P' ?

B is at position P' at t' because

at t B was at position P , and

at t B had the desire to be at P' , and

at t nothing in the world excluded B from being at P'

In the example, the desire (which is usually considered as a kind of future-directed mental state) plays a role similar to that of a potentiality p for being at P' and the explanation for a moving arrow as discussed in Sect. 15.2.2. Indeed this similarity can be traced back in history, for example, to Aristotle:

Now we see that the living creature is moved by intellect, imagination, purpose, wish, and appetite. And all these are reducible to mind and desire (Aristotle, 350 BC b, *De Motu Animalium*, Part 6).

And so what we do without reflection, we do quickly. For when a man actualizes himself in relation to his object either by perceiving, or imagining or conceiving it, what he desires he does at once. For the actualizing of desire is a substitute for inquiry or reflection. I want to drink, says appetite; this is drink, says sense or imagination or mind: straightway I drink. In this way living creatures are impelled to move and to act, and desire is the last or immediate cause of movement, and desire arises after perception or after imagination and conception. And things that desire to act now create and now act under the influence of appetite or impulse or of desire or wish (Aristotle, 350 BC b, *De Motu Animalium*, Part 7).

As an extension of the idea of potentiality he also describes what today is often called means-end reasoning. He explicitly summarizes that ‘things in the soul’ control action:

Now there are three things in the soul which control action and truth - sensation, reason, desire. Of these sensation originates no action; this is plain from the fact that the lower animals have sensation but no share in action (Aristotle, 350 BC c, *Nicomachean Ethics*, Book VI, Part 2).

15.3.2 Can ‘Things of the Soul’ Move Objects?

In the analysis in Sect. 15.3.1 properties of ‘mind and desire’ are mentioned as the source of motion of a living being. Aristotle (350 BC b) shows how the occurrence of certain internal (mental) state (desires) within the living being entail or cause the occurrence of an action in the external world; see also Nussbaum (1978). Such internal state are sometimes called by him ‘things in the soul’, ‘states of character’, or ‘moral states’. In that time such ‘things’ were not considered part of the physical world but of the ghost-like world indicated in this case by ‘soul’, similar to what happened for the case of nonliving entities. So, in this context the explanation that such a creature’s position gets changed is that there is a (ghost-like) state of the soul driving it. How such nonphysical ghost-like states can affect physical states remains unanswered, also for this case of living entities. Over time within Philosophy of Mind this has been felt as a more and more pressing problem. Within Philosophy of Mind nowadays a well known manner to characterise mental state is based on the notion of functional or causal role; e.g., Kim (1996, 1998); see also Chap. 1,

Sect. 1.2. The analysis above illustrates how mental states may have a causal role with respect to future states or behaviour in the physical world, which makes that they can be viewed as specific cases of potentialities for the states or behaviour they cause. The idea that mental states can cause behaviour is called *mental causation* (e.g., Kim 1996). But the problem with this is how exactly can nonphysical mental states cause effects in the physical world? Mental causation can be seen as similar to the idea of ghost-like causation discussed in Sect. 15.2.2 for nonliving entities. It has the same problem of physical effects caused by nonphysical states, without any mechanism known for such an effect.

15.4 Explaining Changed States by Introducing Potentialities

The assumptions discussed in Sects. 15.2 and 15.3 focus on motion of living and nonliving objects and the possibility to include concepts (potentialities) in the ontology to conceptualise states that describe properties of (changed) future states concerning motion. The current section addresses this idea from a more generic perspective addressing any change, and provides some more detail.

15.4.1 *Potentialities and Their Actualisation as a General Perspective on Dynamics*

Aristotle did introduce such a concept; he called it *potentiality* (to move), or *movable*. The difference between the arrow at rest at time t at position P and the snapshot of the moving arrow at t at position P is that at time t the former has no potentiality p to be at P' , whereas the latter has. This explains why at a next instant t' the former arrow is still where it was, at P , while the latter arrow is at a different position P' . Aristotle did not only consider changes of positions (due to locomotion), but also other types of change, for example, a young man becoming an old man, and a cold object becoming hot. For each of these types of changes a specific type of potentiality is considered; e.g., the potentiality to be at position P' , the potentiality (of a young man) to be an old man, the potentiality (of a cold object) to be hot.

In general, if the potentiality p (occurring in a state S) to have state property X has led to a state S' where indeed X holds, then this state property X of state S' is called the *fulfilment* or *actualisation* of the potentiality p for X occurring in state S . He expresses his view on potentialities and their actualisation as follows:

We have now before us the distinctions in the various classes of being between what is full real and what is potential.

Def. The fulfilment of what exists potentially, in so far as it exists potentially, is motion - namely, of what is alterable qua alterable, alteration: of what can be increased and its opposite what can be decreased (there is no common name), increase and decrease: of what can come to be and can pass away, coming to be and passing away: of what can be carried along, locomotion.

The same thing, if it is of a certain kind, can be both potential and fully real, not indeed at the same time or not in the same respect, but e.g. potentially hot and actually cold [from (Aristotle, 350 BC a, Physics), Book III, Part I].

15.4.2 *Derivatives as Potentialities for Variables in Dynamical Systems*

Consider a jump from the time of Aristotle to now. As discussed in Chap. 2 dynamics of continuous process is often described by dynamical systems that involve a number of states or variables X_i with different values $X_i(t)$ for different points in time t . These values $X_i(t)$ for different time points together describe the succession of overall states S for the different time points (for example, see Chap. 2). Changed state properties of a state S' at some time point t' compared to a state S at an earlier time point t have the form that at least one of the values $X_i(t')$ at t' is different from the value $X_i(t)$ at t : at least one of the X_i has changed its value from t to t' .

Now focus on such a change of any continuous variable X , from a time point t to a time point t' , with Δt the time difference $t' - t$. During this Δt the value of X changes from $X(t)$ to $X(t')$. The derivative $\mathbf{d}X/\mathbf{d}t$ of this variable is usually considered the 'rate of change' for X . What exactly does this mean? Together with the value $X(t)$ at time point t , the derivative $\mathbf{d}X(t)/\mathbf{d}t$ at t determines the state of $X(t')$ at $t' = t + \Delta t$ in the following manner. For small Δt as an approximation with $\Delta X(t) = X(t + \Delta t) - X(t)$ it holds

$$\begin{aligned}\Delta X(t)/\Delta t &= \mathbf{d}X(t)/\mathbf{d}t \\ \Delta X(t) &= \mathbf{d}X(t)/\mathbf{d}t \Delta t\end{aligned}$$

This shows how the change $\Delta X(t)$ of the state of X from t to t' is fully determined by the derivative $\mathbf{d}X(t)/\mathbf{d}t$ and Δt . So, the derivative $\mathbf{d}X(t)/\mathbf{d}t$ can be considered a potentiality for this change with the new value $X(t) + \Delta X(t)$ as its actualisation. Note that this can be rewritten into the following format:

$$\begin{aligned}(X(t + \Delta t) - X(t))/\Delta t &= \mathbf{d}X(t)/\mathbf{d}t \\ X(t + \Delta t) &= X(t) + \mathbf{d}X(t)/\mathbf{d}t \Delta t\end{aligned}$$

This again shows how the value $X(t + \Delta t)$ in the next state at $t + \Delta t$ is determined by the values for $X(t)$ and $\mathbf{d}X(t)/\mathbf{d}t$ at t ; in Chap. 2, Sect. 2.8 it was discussed how this basic format can be used for simulation.

The case of Zeno (who did not have the machinery of derivatives as available nowadays) can also be rephrased in these terms. The difference between the moving and non moving arrow can be defined as a difference in speed, which is the derivative of the arrow's position $X(t)$ over time. But the question how this derivative $\mathbf{d}X(t)/\mathbf{d}t$ at t can be considered a state property at t in general still may not be easy. Is this not a sneaky manner to let a notion of motion enter again through the back door, while Zeno's analysis (see Sect. 15.2.1) made it clear that there is not such a thing as a state property? This will be discussed below in Sect. 15.6.1.

15.4.3 *What Kind of State Properties Are Potentialities?*

The similarity in explanatory pattern for different cases of dynamics in different domains leads to the question what potentialities actually are in these different cases and domains. For example, Zeno claimed that a moving arrow at t does not differ in state from an arrow at rest at the same position. Even relating this to a concept as velocity which does make a difference at t does not fully solve this issue, as will be shown in more detail in Sect. 15.6.1 below. It may seem strange to attribute such invisible ghost-like state properties to certain (living or nonliving) objects. Wouldn't the use of such vague concepts stand in the way of a genuine physical description of the world? Recall, however, that adding the concept potentiality to the state ontology was done to solve an explanatory problem that otherwise was hard to solve: how to explain that two given arrows in exactly the same position, one arrow is in another position in a next state, whereas the other arrow still is in the same position. So, simply banning such an unclear concept leaves us with this problem, which actually can be viewed as a problem of non-determinism: without such an additional property the state at t does not fully determine the state at $t' > t$.

There is a longstanding discussion in Philosophy of Mind on the existence and place of mental state, and the problem of mental causation: how can ghost-like mental state ('things of the soul') make physical things (an organism) move; also see Kim (1996) and Sect. 15.3.2. Considered at a more general and abstract level, this has much in common with the discussion in Sect. 15.2 on the existence and place of potentialities for nonliving entities: in how far are such potentialities real? Assuming potentialities as additional state properties, just to make this difference and in this way guarantee determinism would be a cheap and artificial solution (just defining the problem away) if it is not shown how potentialities can obtain a solid place as genuine state properties. Moreover, their presence is not very useful for actual predictions if there is no manner by which it is possible to find out or predict whether and in which form they are there. This issue will be discussed in more detail in Sect. 15.6, after some further discussion about potentialities in Physics in Sect. 15.5.

15.4.4 *Summary of Assumptions Underlying Potentialities*

As a unifying perspective the following criteria will be considered as characteristic for a perspective based on anticipatory state properties or potentialities:

- *succession of states*
The world occurs in successive states at different points in time.
- *state-based ontology*
Within an explanation or description only states and their properties are used, based on a state ontology. In particular, no concepts for actions, events, transitions between states, or processes are used.
- *necessity of anticipatory state properties*
For each specific state property *a* that occurs in a given state there exists a specific anticipatory state property *p* related to property *a* that occurs in a preceding state, in conjunction with some additional conditions on specific circumstances in this state (i.e., no obstruction occurs of the actualisation of *a* by *p*).
- *sufficiency of anticipatory state properties*
If *p* is a specific anticipatory state property related to the occurrence of a specific state property *a*, and property *p* occurs in a state, then, given suitable further circumstances (i.e., no circumstances obstructing the actualisation of *p*), in a subsequent state property *a* will actually occur.
- *state-based or temporal grounding of anticipatory state properties*
For each anticipatory state property *p* there is a specific characteristic either in the past and/or in the current states that guarantees the occurrence of *p*.

15.5 Potentialities in Physics

In later times successors of Aristotle, such as René Descartes (1596–1650), Christiaan Huygens (1629–1695), Isaac Newton (1643–1727) and Gottfried Wilhelm Leibniz (1646–1716), among others, have addressed the question how to further develop the phenomenon of change or dynamics and, in particular, the concept potentiality within Physics. Some contributions of these will be discussed in this section. Indeed they succeeded in giving certain types of potentialities a well-respected place in modern physics (actually in (1644, 1998) more than one way).

To obtain a better understanding of the concept of potentiality (which he called *quantity of motion*, or *tendency to motion*), Descartes (1644, 1998) did some reflection on objects of different sizes.

Now, although this motion in moved matter is nothing other than its mode, nevertheless it has a certain and determinate quantity, which we easily understand to be able to be always the same in the whole universe of things, even though it be changed in its individual parts. So it is evident, as we think, that when one part of matter is moved twice as fast as another, and

this second [part of matter] is twice as large as the first, there is as much motion in the smaller as in the larger ... (Descartes 1644, Principles of Philosophy, Part II, Paragraph 36).

Descartes took the product mv of mass m and velocity v of an object as an appropriate foundation for its potentiality to be in a changed position, or quantity of motion. Thus he related the vague concept potentiality to other, better known concepts. Notice that this anticipatory state property 'quantity of motion' is a relative potentiality: the actualisation of a given quantity of motion entails being at another position as specified by this quantity relative to the current position, and not as being at some absolutely specified position. In modern physics this 'quantity of motion' concept is called *linear momentum*, or just *momentum*, and the conservation, for example, during elastic collisions, is called the 'law of momentum conservation'. Newton incorporated this notion in his approach to motion; actually the law of momentum conservation as formulated by Descartes has a strong relationship to Newton's second and third law. This is one way in which a concept 'potentiality' was given a well-respected place in Physics, in particular in classical mechanics.

Huygens (1629–1695), and later his student Leibniz (1686a, 1686b, 1989, 1956, 1991), used a different way to give a concept 'potentiality' a place in Physics. Leibniz called this concept *vis viva* (*living force*), or *motive force*, or *moving force*, or *force of motion*, or *power*. By incorporating results from experimental work of Galileo, Leibniz has shown that his notion motive force and Descartes' notion quantity of motion are different concepts:

Thus, through the resolution of bodies into parts, the speed, or space and time, being conserved, we had inferred, demonstrated, that given the same speeds the powers were proportional to the bodies. Similarly, we have demonstrated, which is paradoxical, but absolutely true, that, the body being conserved, time and space being resolved jointly (for otherwise the case given could not be divided in several cases congruent with each other while different), given the same bodies, the powers are proportional to the square of speeds (Leibniz 1991, II, §E, p. 816).

So Leibniz claimed that the potentiality 'motive force' was proportional not with the velocity v as in the case of 'quantity of motion', but with the square of the velocity. In this way Leibniz put the foundation for the law of conservation of energy, in this case involving kinetic energy (which actually was later taken $\frac{1}{2}mv^2$) and potential energy, and exchange between the two. So, within the development of Physics, the potentiality of an object to be at a different position was differentiated in two forms: linear momentum (Descartes' quantity of motion) and kinetic energy (Leibniz's motive force). Today both still are part of Physics; both were expressed in terms of mass m and velocity v of the object (mv resp. $\frac{1}{2}mv^2$).

In a broader sense Leibniz aimed at developing what he called a science of power and action, or a *science of dynamics*; in the Specimen præliminare of the Dynamica he states:

I judged that it was worth the trouble to muster the force of my reasonings through demonstrations of the greatest evidence, so that, little by little, I might lay the foundations for the true elements of the new science of power and action, which one might call

dynamics. I have gathered certain preliminaries of this science for special treatment, and I wanted to select a ready specimen from these in order to excite clever minds to seek truth and to receive the genuine laws of nature, in place of imaginary ones (GM VI, p. 187; Leibniz 1989, p. 107).

Note that to relate a notion of potentiality to other state properties, both Descartes and Leibniz made use of mass and velocity as state properties. This provides a more operational and practically useful type of potentialities: they can be calculated by the formulae mv and $\frac{1}{2}mv^2$ as soon as mass and velocity can be determined. However, in how far velocity, and more in general any derivative of a state property, can be considered a genuine state property is a not so easy question, which will be discussed in Sect. 15.6. It might be the case that the formulae discussed above just express some potentialities in another potentiality, namely velocity. At least an advantage is that the focus of the analysis can be directed to velocity, which will be addressed in particular in Sect. 15.6.1.

15.6 What Kind of Property Is a Potentiality: Getting Rid of Ghosts?

Within Philosophy of Mind it has been described how mental states such as desires can play a causal role, thus functioning as potentialities for behaviour. Playing a causal role means that they are caused by some other states which may include other mental states and states involved in sensing the world, and in turn they cause other states which may include other mental states and actions or behaviour of the organism in the physical world. However, especially the latter type of causation, also called mental causation was felt as a difficult problem: how can mental states make things move in the physical world, if they are nonphysical, ghost-like state properties? Such interaction from ‘things of the soul’ to the physical world is considered problematic, as usually it is assumed that physical effects have physical causes, and no mechanism is known for interaction from nonphysical entities to physical entities. Given the role of potentialities in Physics as described in the previous section, could something be learned from that?

Within Physics, potentialities have found their place in different manners. Basic concepts such as momentum, kinetic energy, and also force can be considered variants of potentialities. In Sect. 15.5 momentum and kinetic energy were discussed; the concept force will be discussed in Sect. 15.10 in the context of higher-order potentialities and the exchange of potentialities by interaction. Both for momentum and kinetic energy a conservation law has been found, and both concepts have been expressed in terms of mass and velocity (see Sect. 15.5). Does this mean that in these two forms, potentialities have become genuine state properties, because they are definable in terms of other genuine state properties? Even leaving relativity theory aside, this is not a simple question. A straightforward answer would be: indeed, potentialities are genuine state properties because they are

defined in terms of mass and velocity which are assumed to be genuine state properties. For the sake of simplicity accepting this claim for mass, a question, however, remains what kind of state property velocity $V(t)$ at time t is.

15.6.1 *Why Velocities and Derivatives by Themselves Are not Genuine State Properties*

What type of state property is velocity? Could it be the case that velocities, and more in general derivatives of continuous variables, are having a ghost-like status? This question will be addressed first, by focussing on velocity $V(t)$; but the same analysis applies to the derivative $dX(t)/dt$ of any variable X .

A first approach is to take velocity $V(t)$ to be distance traversed divided by time passed over some chosen time interval from t'' to t ; i.e.:

$$V(t) = (X(t) - X(t'')) / (t - t'')$$

with $X(t)$ distance traversed at t and $X(t'')$ distance traversed at t'' . This definition involves states at different points in time t and t'' , so it is not based on one state at one time point. This notion of velocity actually is velocity over the given time interval from t'' to t , so a property of a sequence of states indexed by the time points of the interval, or, to simplify it a bit, a property of a pair of states for the starting point and the end point of the time interval. This is not what one would call a genuine state property for the state at time t .

For a second trial, consider a modern variant of the moving arrow discussed in Sect. 15.2. Suppose there are two snapshots of a car at time t : one was driving (with 50 km/h) at t and one is in rest. In this snapshot of the driving car everything is frozen in the position at t . So everything looks the same as the car in rest, the car itself, and its wheels, for example. But wait, within the interior of the driving car there is one visible difference! The *speedometer* of the frozen car indicates 50 km/h, whereas in the car in rest it indicates 0 km/h. So does this finally provide a real and observable difference between the physical states of the two cars? Should velocity at some point in time be identified with what this speedometer displays? Indeed, at a point in time t the position of the pointer of a speedometer, or the number displayed is a genuine state property. Would this offer an appropriate solution to get rid of the ghost status of velocity? A first general objection may be that this state property is just the position of the pointer, not velocity. For every type of object and speedometer a different state concept would arise: think of speedometers for cars compared to those of airplanes, ships, rockets; and what about the velocity of a bird or an approaching meteor, should they also have speedometers? Even if for a certain class of objects, such as cars, a standardisation would be reached for a speedometer, then still the position of the pointer of the speedometer is itself not velocity; at most it has a *relation* to velocity. What kind of relation is this?

Looked at it in some more detail, the position of the pointer of a speedometer results from or is affected by the actual motion, so there is a small time delay between having some velocity and what the pointer displays. Therefore the pointer actually indicates speed at time points $t'' < t$, which, although close to t are not exactly equal to t . Thus the pointer position, which is a state property of the state at t , actually relates to velocity in states at $t'' < t$, not to a state property of the state at t . This makes clear that the speedometer concept will not help much to make velocity a genuine state property of the state at time t . Finally, also from another angle it is quite difficult to imagine that this pointer position can be seen as a potentiality, as a potentiality is supposed to cause the changed position after t . And certainly the pointer position is not moving the car: the pointer position itself does not affect anything but the image in our eyes and certainly not the (changed) position of the car at the next instant; velocity does affect this position. Apparently the problem is not solved by considering the pointer position as a candidate for potentiality.

A third possible approach is what is sometimes called the notion of *instantaneous velocity*. In modern physics and mathematics, for continuous processes that satisfy sufficiently strong conditions of smoothness, this is usually defined as a limit:

$$V(t) = \lim_{t'' \rightarrow t} (X(t'') - X(t)) / (t'' - t)$$

Note that this limit is defined in terms of the whole family or sequence of states around t , i.e., in terms of the state properties $x(t'')$ for all t'' in a neighbourhood of t . In mathematical terms this limit can be defined as

$$\forall \varepsilon > 0 \exists \delta > 0 \quad \forall t'' [0 < |t'' - t| < \delta \Rightarrow |(X(t'') - X(t)) / (t'' - t) - V(t)| < \varepsilon]$$

It is clear that this statement refers to a whole sequence of states for time points t'' around t . In this sense also the notion of instantaneous velocity does not provide a good solution for a foundation of the potentiality velocity as a genuine state property.

Note that the approaches discussed above at least provide some practically useful methods to determine an approximate value for the velocity as a potentiality. The first approach can be used by just taking some t'' just before t and then make the calculation of the velocity over the interval from t'' to t . This may not be the best approximation for velocity at t , but at least gives some value. The second approach is just by using a measuring device (speedometer). This also gives an approximate value, due to possible inaccuracies in the measuring device. The third method could provide the best approximation, but also not a perfect value, as such a limit can only be approximated by a numerical method up to a certain accuracy.

The analysis of the notion of velocity shown above can also be done in a similar manner for the derivative dX/dt of any variable X . Recall from Sect. 15.4.2 that

together with the value $X(t)$ at time point t , the derivative $\mathbf{d}X(t)/\mathbf{d}t$ at t determines the state of $X(t')$ at $t' = t + \Delta t$. For small Δt as an approximation it holds

$$X(t + \Delta t) = X(t) + \mathbf{d}X(t)/\mathbf{d}t \Delta t$$

This shows how the value $X(t + \Delta t)$ in the next state at $t + \Delta t$ is fully determined by the values for $X(t)$ and $\mathbf{d}X(t)/\mathbf{d}t$ at t . The issue that has been discussed above for the case of velocity in principle applies to any derivative $\mathbf{d}X(t)/\mathbf{d}t$ at t , so it may not be easy to find out how this $\mathbf{d}X(t)/\mathbf{d}t$ can be considered to be a genuine state property at t .

15.6.2 *Ghost-like Properties or Temporal Relations Involving Genuine Properties?*

In summary, it turns out that to define a potentiality p for a state at a time point t , states at different time points may have to be taken into account, not only the one state at t . So, do we have to admit that addressing motion and change by extending the state ontology by some form of additional state properties for potentialities is failing? The answer on this question seems to be: yes and no. The answer is 'yes' in the sense that in the three possibilities considered here, there has not been found anything physically real in the state as a basis for a potentiality p as a genuine state property. The answer is 'no' in the sense that the historical developments as discussed in this section have provided quite powerful mathematical means (calculus, differential equations) to model all kinds of problems in diverse application areas. In our daily life we all rely on artefacts constructed using classical mechanics; e.g., bridges, buildings, transportation means. Given that this conceptual machinery works quite well in predictions, makes that the question is still there: what is it that makes this machinery so successful? You would not like to answer this question by: ghosts!

From a general perspective, avoiding the use of a concept for motion itself as being too ghost-like, the basic ontology for state properties apparently is insufficient for explanations of dynamics. Given this problem, in principle two ways out are possible: *either* (1) extend the assumed state ontology and state properties expressed in terms of them by additional anticipatory state properties or potentialities p , to be able to discriminate states that are at rest and states that are going to change, *or* (2) keep the basic state ontology and state properties the same but extend the states that can be used in such an explanation from the current state to states in the past in addition. So, more specifically, the following two ways can be pursued to solve the problem:

- (1) *Extending the state ontology by introducing anticipatory state properties*
 Assume that to conceptualise a state, the state ontology has more ontological elements and state properties than only the apparent ontology and state properties. In particular, a state can be conceptualised using an additional type of *anticipatory state property* p in the state ontology: the *potentiality* to subsequently get different state properties, i.e., the potentiality for the state to become changed.
- (2) *Exploiting temporal relationships: involving states over different time points*
 Explaining why a state at time t' has a different state property is not possible on the basis of one state at $t < t'$, but needs to take into account a *history of different previous states at times t'' before t'* . In Kim (1996, pp. 200–202) a notion of (temporal) relational specification for mental states is discussed that could play a role in this perspective. Also the three methods discussed above to determine velocity make use of such a history of states before t .

Above it was discussed that option (1) does not provide a satisfactory solution for the case of velocity, and derivatives more in general. Choosing for this option makes that one still is left with ghost-like state properties that cannot be related to physical reality. So, better go for the second option? The different trials for velocity above all ended up in relations with states at other time points than t , so that would suggest that option (2) is the option to choose. However, also option (2) has some problems. One severe problem is that if velocity or derivative is defined in terms of states at t'' before t , option (2) violates the state-determined system assumption. The state at t' does not only depend on the state at t , but via the temporal relations also on states at t'' before t . As state-determined systems are the focus here, this is an unforgivable shortcoming of option (2).

So for a good foundation of state-determined systems only option (1) remains, with its problems. One possible solution for (1) is the *predictable ghost option*: accept that a potentiality is not a genuine state property, so it is still ghost-like, but by using the temporal relations as in (2) (see also Sect. 15.6.1), at least it can be predicted from previous states before t when it will occur. This is how the potentialities momentum and kinetic energy and the notion velocity are handled in Physics, with success. Adding these potentialities as artificial state properties at time t is not that harmful, as they are fully determined by genuine state properties in previous states before t : they may occur like ghosts, but at least they are predictable ghosts. In this way still a state-determined system is obtained, although this was achieved in an artificial manner.

To address the problems of option (1) there is still some other possibility that may be considered as a sort of solution for these problems: the *realisation option*; this will be discussed in Sect. 15.8. As a next step, to get some inspiration for this type of solution, in Sect. 15.7 first the role of potentialities in some approaches in Artificial Intelligence and Computer Science is discussed: modeling based on causal relations or transition systems.

15.7 Potentialities for Causal Relations and Transition Systems

An often used method (in Artificial Intelligence, Computer Science and related areas) to specify how a state in a system may change is known as transition systems; causal relations can also be described in this setting.

15.7.1 Transition Systems and Causal Relations

Transition systems are collections of specifications that each consist of a pair (φ, ψ) , also denoted as $\varphi \rightarrow \psi$ and sometimes called a *transition rule* with antecedent φ and consequent ψ . In this specification:

- the first description φ indicates a combination of state properties for the current state (for example a conjunction of basic state properties, or a disjunction, or a disjunction of conjunctions)
- the second description ψ indicates a state property for the next state.

The idea is that if the combination of properties specified in the first description holds in a (current) state, then in a next state the properties specified by the second description will hold. In fact causal relations can be considered a specific case of this, in which φ causes ψ .

This approach is illustrated by a simplified model of traffic lights at a crossing of two roads A and B, where traffic on A has priority over traffic on B. For example, if no approaching traffic is sensed on road A, then the traffic light for road B is set green, and for road A red. Such a simple scenario can be described in transition system format as follows:

$$\begin{aligned} \text{traffic_on_road_A} &\rightarrow \text{green_light_for_road_A} \\ \text{traffic_on_road_A} &\rightarrow \text{red_light_for_road_B} \\ \text{no_traffic_on_road_A} &\rightarrow \text{red_light_for_road_A} \\ \text{no_traffic_on_road_A} &\rightarrow \text{green_light_for_road_B} \end{aligned}$$

These relations can be interpreted as well as a set of causal relations. For example, the presence of traffic on road A causes the light for road B to be red.

Based on such a specification a trace of subsequent states is made (in an iterated parallel fashion) as follows:

- Given a current state S , take the transition rules for which the antecedent holds in the current state. This is the set of applicable rules.
- Collect the consequents of all applicable rules and obtain the next state S' by modifying S so that all these consequents hold in S' (and the rest of S is persisting).

15.7.2 *Potentialities for Transition Systems and Causal Relations*

How can such a dynamical system model be interpreted in terms of potentialities? For example, consider a state S with no traffic on road A. Then by transition rule

$$\text{no_traffic_on_road_A} \rightarrow \text{green_light_for_road_B}$$

in the next state S' the property `green_light_for_road_B` holds. Therefore in state S the potentiality for `green_light_for_road_B` has to be present, i.e., the state property $p(\text{green_light_for_road_B})$ occurs in state S . Similarly the other transition rules can be interpreted as indications of which potentialities occur in a given state. In general, according to this interpretation a transition system specifies for each state which potentialities occur: for each transition rule $\varphi \rightarrow \psi$, if in a state S its antecedent φ holds, then in this state S also the potentiality $p(\psi)$ for ψ occurs. Thus a transition rule $\varphi \rightarrow \psi$ can be interpreted as an implication

$$\varphi \rightarrow p(\psi) \quad \text{within any state } S$$

describing a relationship between state properties in the given state S . If in the transition rule $\varphi \rightarrow \psi$ the antecedent φ incorporates all antecedents that lead to consequent ψ (for example, φ is the disjunction of all of them) it is even possible to have a bidirectional implication:

$$\varphi \leftrightarrow p(\psi) \quad \text{within any state } S$$

In a more general setting, suppose in a temporal-causal network states X_i for $i = 1, \dots, k$ have impact $\omega_{X_i, Y}$ $X_i(t)$ on state Y , and $c_Y(\dots)$ is the combination function for these impacts. Then the combined impact on Y is $c_Y(\omega_{X_1, Y} X_1(t), \dots, \omega_{X_k, Y} X_k(t))$. This combined impact as an aggregation of the single impacts can be considered an aggregated state X with value $V = c_Y(\omega_{X_1, Y} X_1(t), \dots, \omega_{X_k, Y} X_k(t))$ which is expressed in terms of the states X_i . This state X with its value V can be related to the potentiality $p(\psi)$ by $\varphi \leftrightarrow p(\psi)$, where ψ stands for $Y(t) = V$ and φ stands for $X(t) = V$.

This relation $\varphi \leftrightarrow p(\psi)$ between φ and $p(\psi)$ within any state S suggests a solution of the problem how to get rid of the ghost-like status of potentialities: by relating them to other state properties in the same state. By directly relating the (by itself ghost-like) state property $p(\psi)$ to a state property φ , this gives $p(\psi)$ a more genuine status in the state. Moreover, the effect of $p(\psi)$ on ψ can now be explained as a causal effect of φ on ψ , which is no ghost-effect anymore, and is practically useful to actually determine the next state S' (as long as causal effects are assumed real). This type of solution for the status problem of potentialities will be discussed in more detail and in a more general setting in Sect. 15.8.

15.8 Realisers for Potentialities and the Role of Differential Equations

A not yet fully considered possibility to get rid of the ghost-like character of potentialities is by embedding them better in the states in which they occur. This can be done by trying to identify relationships between a potentiality and other, more real properties of the states in which they occur. Such a state property, making the potentiality more real, is called a *realiser* of the potentiality. First it is discussed how this idea was developed within Philosophy of Mind.

15.8.1 *Realisers of Mental States in Philosophy of Mind*

For mental states this is an approach that has been studied extensively in the literature on Philosophy of Mind, in order to get rid of the ghost-like character of the mind, and in particular of mental state; e.g., see Kim (1996). For example, this provides a solution for the well-known problem of *mental causation*: how can mental states have effect on the physical world?

In this solution a mental state M is related to a neurological or physical state R , called a *realiser* of M , that always co-occurs with M ; for every point in time mental state M occurs in the state at t if and only if R occurs in the state at t :

at each t for any state at t it holds $M \leftrightarrow R$

Given such a *realisation relation*, a solution for the problem of mental causation can be obtained as follows. Suppose M is a mental state that is assumed to cause a physical action effect A . A causal relation $M \rightarrow A$ would be seen as a ghost-like state property ('thing of the soul') causally affecting a physical property, which is considered problematic. However, now there is a way out of this. When an effect of M on A is considered, in reality this can be considered an effect of the physical realiser R of M on A : a physical causal relation $R \rightarrow A$. The latter relation is just a relation between two properties in the physical domain, so there is nothing ghost-like remaining. A lot more can be said about this notion of *reduction* or *realisation relation* (e.g., for more details see Kim 1996), for example, about whether for the human or animal mind such realisers always exist and if so, if there may exist more of them (multi-realisation). This will not be discussed further at this point, but the idea itself is adopted here for potentialities, beyond the 'mind' context, thereby for the sake of simplicity not considering multi-realisation for now; however, see Sect. 15.11 for some more details on multiple realisation.

15.8.2 Realisers of Potentialities from a More General Perspective

Now compare the analysis in Sect. 15.8.1 to the analysis of causal relations and transition systems in Sect. 15.7.2. There a causal relation $\varphi \rightarrow \psi$ was considered, and this was related to a potentiality $p(\psi)$ for ψ by a relation $\varphi \leftrightarrow p(\psi)$ within a state. This has a clear similarity with a realisation relation $M \leftrightarrow R$ within Philosophy of Mind, where mental state property M corresponds to potentiality $p(\psi)$ and the physical realiser R corresponds to φ .

For the general case, inspired by these cases considered above, if in all states a potentiality p always co-occurs with a certain genuine state property c (which can also be a combination of more basic state properties), such a co-occurring property c is called a *realiser* for p : for every point in time t , state property p occurs in the state at t if and only if c occurs in the state at t :

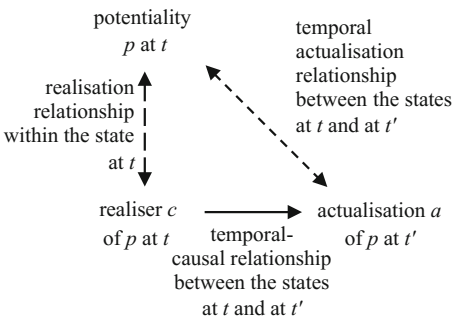
$$\text{at each } t \text{ for any state } S \text{ it holds } p \leftrightarrow c$$

See also Fig. 15.2; the vertical bidirectional arrow indicates the *realisation relation* between p and c .

If such a realiser exists, it is possible to get rid of the problem that a ghost-like potentiality given by state property p has a causal effect on its actualisation a , in a similar manner as the solution of the problem of mental causation in Philosophy of Mind. The idea is that it is not p itself which has this causal effect, but it is its realiser c , via the horizontal arrow in Fig. 15.2. This relation $c \rightarrow a$ can be just a causal relation between genuine state properties, so nothing ghost-like there.

However, there are cases in which it no realiser exists at all. For example, this happens for a freely moving object in space, where a potentiality (momentum or velocity) occurs that is independent of the other properties of the present state. In this case the potentiality at time t depends on the history before t and is independent of properties of the world state at t . To be able to cover such motion effects, it seems that the only way to relate potentialities to other state properties is by relating them

Fig. 15.2 Realisation relation $p \leftrightarrow c$ for a potentiality p for a



to properties of states at time points different from t (i.e., by using temporal relationships). But there are many other cases for which this approach based on realisation relations in the same state at t still can work.

15.8.3 Realisers for Derivatives: First-Order Differential Equations

As a next step it is shown how smooth continuous state-determined systems can be described by realisation relations. In Sect. 15.6 it was discussed how the derivative $\mathbf{d}X/\mathbf{d}t$ of a continuous variable X at a certain time point can be viewed as a potentiality which determines the next state at time $t + \Delta t$ of that variable. In this section it is discussed in which form a realisation relation of such a potentiality occurs, and more in particular, how differential equations play a role in this. Let p_X be such a potentiality (i.e., change rate) for variable X , i.e., $p_X = \mathbf{d}X/\mathbf{d}t$. How can this potentiality be related to other state properties? As a special case, the relationship of (the value of) p_X to other state properties can focus on properties that can be expressed in terms of (the value of) X . A plain case of this idea is when a value V of p_X in a state is considered always to co-occur with this value V for some expression for a function F in the value of X in the same state:

$$p_X(t) = V \leftrightarrow F(X(t)) = V$$

This shows a bi-conditional form for the co-occurrence of the two properties in a state at t , where the right hand side of the ‘if and only if’ is the realiser of the potentiality at the left hand side. An alternative way to express the same biconditional relationship is:

$$p_X(t) = F(X(t))$$

Keeping in mind that the potentiality p_X is the derivative $\mathbf{d}X/\mathbf{d}t$ of X , the latter way of expressing can be also written as

$$\mathbf{d}X(t)/\mathbf{d}t = F(X(t))$$

This expression is the usual notation for a *first-order differential equation*. So, this differential equation allows to relate the potentiality $p_X(t) = \mathbf{d}X(t)/\mathbf{d}t$ at time t to other state properties of the state at t . As an example, take the function F defined by:

$$F(X) = \alpha X(1 - X)$$

For this example, potentiality $p_X(t)$ is related to another state property at time t as follows:

$$p_X(t) = \alpha X(t) (1 - X(t))$$

In the usual notation for a differential equation this is also formulated as

$$\mathbf{d}X/\mathbf{d}t = \alpha X(t) (1 - X(t))$$

It turns out that first-order differential equations can be understood from the conceptual framework based on potentialities as realisation relations for potentialities. The differential equation format

$$\mathbf{d}X/\mathbf{d}t = F(X(t))$$

expresses in a variety of cases how a potentiality p_X relates to another state property. Moreover, this can easily be extended to a system of multiple differential equations for multiple states, such as

$$\begin{aligned} p_X(t) &= \mathbf{d}X/\mathbf{d}t = F(X(t), Y(t)) \\ p_Y(t) &= \mathbf{d}Y/\mathbf{d}t = G(X(t), Y(t)) \end{aligned}$$

where each of the potentialities $p_X = \mathbf{d}X/\mathbf{d}t$ and $p_Y = \mathbf{d}Y/\mathbf{d}t$ has a realisation relation to a combination of the state properties X and Y , defined by F and G , respectively:

In a discretised form a (*first-order*) *difference equation* can be considered:

$$\begin{aligned} \Delta X/\Delta t &= F(X(t)) && \text{with } \Delta X = X(t') - X(t) \text{ and } \Delta t = t' - t \\ (X(t') - X(t))/\Delta t &= F(X(t)) \\ X(t') - X(t) &= F(X(t))\Delta t \\ X(t + \Delta t) - X(t) &= F(X(t))\Delta t && \text{with } \Delta t = t' - t \\ X(t + \Delta t) &= X(t) + F(X(t))\Delta t \end{aligned}$$

Given that $F(X(t))$ is a realizer for potentiality $p_X(t)$, the last line can be re-interpreted as the standard pattern for a potentiality for a continuous state X , describing how the next state at $t + \Delta t$ is determined by the current state at t :

$$X(t + \Delta t) = X(t) + p_X(t) \Delta t$$

So, consider the question: why do (first-order) differential equations or difference equations exist? In principle, this is a difficult question, but in the light of the above it could be answered by: in order to get rid of ghosts! Of course, this answer does only indicate a purpose of differential equations once they are available (what is

their use), not the fundamental question why it is possible to find them in so many disciplines and application domains. That question is more difficult to answer. Maybe their availability should be considered as just a fortunate opportunity or gift offered by the world as a kind of miracle, revealed initially by Newton and Leibniz and developed further by many followers.

Recall from Chap. 2, Sect. 2.9 that a smooth continuous dynamical system is state-determined if and only if it can be described by a set of first-order differential equations. This means that the above mentioned miracle concerning the existence of first differential equations realising potentialities for the dynamics of some part of the world is equivalent to the existence of a state-determined system to describe the dynamics of that part of the world. In other words, the state-determined system assumption for the dynamics of some part of the world is equivalent to the existence of realisers of the potentialities involved in the dynamics of that part of the world, and as they are equivalent to state-determined systems (see Chap. 2, Sect. 2.9.3), this applies as well to the temporal-causal networks considered in this book.

Note that there are cases for which first-order differential equations do not exist, one example being moving objects such as Zeno's arrows. However, for such cases it often turns out that still higher-order differential equations can be found that provide realisers of higher order potentialities. This will be explained in Sects. 15.9 and 15.10. For example, for the case of moving objects it turns out that although first-order potentialities have no realisers, still second-order potentialities do have realisers. So, the more general form of the mentioned miracle concerns the existence of differential equations of arbitrary orders.

15.9 How to Explain Changed Potentialities

The effect of a potentiality on a future state can be described by relating the present state to the future state. This specification can be viewed as the definition of what it is a potentiality for. A further question is how to specify when (under which past and present circumstances) a potentiality occurs. Could a potentiality be a really ghost-like property for which nobody knows when it will occur? Such unpredictable ghost-like state properties would be the worst. In that case it is not possible to predict changes. How can this be avoided? If some ghost-like state property has to occur, at least let it be predictable.

For the case of empty space, where an object is assumed to have no interaction with other objects, a potentiality may be present because it was present at an earlier point in time and persisted until t . However, it still remains a question how this potentiality was generated in the past: if the potentiality in a new state is different from the earlier one, a question becomes why this is so. This leads to the question addressed in this section of how a *changed potentiality* can be explained.

15.9.1 Introducing Higher-Order Potentialities: Potentialities for Potentialities

The use of higher-order potentialities is one answer to the question where changed potentialities come from. The idea behind higher-order potentialities is simple. To obtain an explanation of changed state properties over time, potentialities were introduced. Potentialities are also changing over time. If they are genuine state properties themselves, it would be reasonable to treat them just like any other state property that changes over time, so why not use this as a general principle? This means that for a potentiality $p^{(1)}$ a so-called *second-order potentiality* $p^{(2)}$ is introduced to explain why $p^{(1)}$ may become changed over time. And of course this process can be repeated for $p^{(2)}$, and so on. This leads to an infinite sequence of *higher-order potentialities*,

$$p^{(1)}, p^{(2)}, p^{(3)}, p^{(4)}, \dots$$

where for each natural number n the potentiality $p^{(n)}$ is called an *n-th-order potentiality*. The idea is the following:

- for a certain point in time t_0 the occurrence of a state property can be determined on the basis of the state at a previous time point $t_1 < t_0$ and, in particular, the first-order potentiality at that time point t_1 .
- the first-order potentiality at t_1 can be determined by the state at a time point $t_2 < t_1$ and, in particular the second-order potentiality at t_2 .
- and so on.

This process can be visualised as depicted in Fig. 15.3.

This shows how the concept of potentiality to explain change of a certain basic state property a can take the form of a single entity, for example one number, to

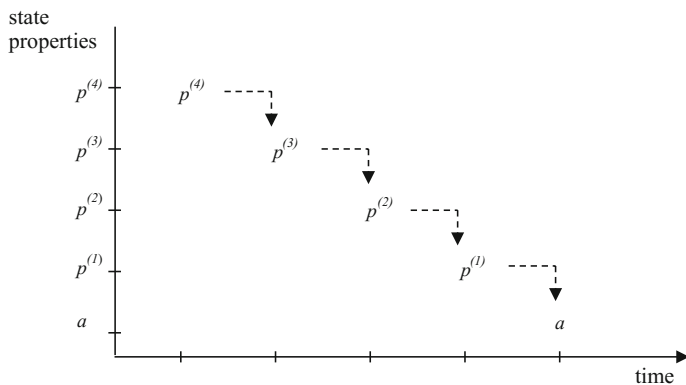


Fig. 15.3 Dynamics based on higher order potentialities

indicate what a changed property in an immediate subsequent state will be, but this can be extended by a large or even infinite number of other (higher-order) entities that can explain changed basic state properties *a* in further future states. Some more details of higher order potentialities can be found in Bosse and Treur (2007, 2008).

15.9.2 Higher-Order Potentialities in Cognitive Models

In Sect. 15.3.1 the concept of desire was interpreted as a (first-order) potentiality. For example, a desire for healthy food makes one eat such food. But where does such a desire itself come from? The general pattern described above is that it comes from a second-order potentiality, that can be described as the desire to have a desire, in the literature sometimes called a *second-order desire* (e.g., Frankfurt 1971). For example, not having a desire for healthy food, someone can still have the desire to have this desire for healthy food. In the literature an interesting discussion can be found about such second-order desires, and the question in how far someone is able and free to have or to choose for a certain desire, or is able to intentionally change a desire (e.g., Frankfurt 1971). Or are second-order desires destined to only stay desires forever and never be actualised? There seems to be no general answer to the question where desires come from and how they can be changed. There may indeed be some form of (ghost-like?) second-order potentiality involved, which maybe itself could be related to observations or beliefs, or in a temporal sense could be based on a history of experiences.

15.9.3 Mathematical Formalisation of Higher-Order Potentialities in Calculus

Strange as the idea of a possibly infinite number of higher-order potentialities may seem at first sight, in a mathematical context (in particular in calculus) this has been worked out quite well. For the discrete case, the idea of difference tables for functions has been developed; see Table 15.1. These differences play the role of

Table 15.1 Dynamics based on a higher-order difference table

Time point	f value	1st-order difference	2nd-order difference	3rd-order difference	4th order difference
0	3	1	1	−3	1
1	4	2	−2	−2	
2	6	0	−4		
3	6	−4			
4	2				

relative potentialities: they indicate the next value not in an absolute sense, but in comparison to the current value.

Such a discrete table may still be finite. However, for the continuous approach higher-order potentialities have been formalised within Mathematics in the form of an infinite number of higher-order derivatives $f^{(k)}(t)$ for $k = 1, 2, 3, \dots$ of a function $f(t)$. The well-known Taylor approximation and Taylor series for sufficiently smooth functions (infinitely often differentiable) show how changes of the value from t to t' (within some given neighbourhood of t) depend on all higher-order derivatives in the form of a convergent infinite summation:

$$f(t') = f(t) + \sum_{k=1}^{\infty} f^{(k)}(t)(t'-t)^k / k!$$

or

$$f(t + \Delta t) = f(t) + \sum_{k=1}^{\infty} f^{(k)}(t)(\Delta t)^k / k!$$

This expression shows how the combination of all (infinitely many) higher-order potentialities, all at t , determines the changed state at the future time points t' .

15.9.4 How to Get Rid of an Infinite Chain of Higher Order Potentialities by Realisers

The analysis above places the question of how to interpret a potentiality as a genuine state property in a different light. Apparently, in the continuous case a potentiality may take the form of a kind of infinitary property, an infinite-dimensional vector of higher-order potentialities; such infinitary properties are far remote from what usually are understood as genuine state properties. In some cases maybe only changes that involve a finite number of higher-order potentialities have to be considered. For example, within a constant gravitation field, the second-order potentiality (the acceleration, which is the second-order derivative of the distance) is constant (9.8 m/s^2), and hence no third- or higher-order potentiality is needed: they are all zero; this will be discussed in some more detail in Sect. 15.10. Note, however, that further away in the solar system or the universe, if an object is approaching the earth, gravitation will increase over time, so this assumption of constancy will not always be fulfilled.

In general a possibility to get rid of the infinite chain is obtained when for some n the n -th order potentiality $p^{(n)}$ has a realiser c :

$$\text{at each } t \text{ for any state at } t \text{ it holds } p^{(n)} \leftrightarrow c$$

See Fig. 15.3 for a case in which $n = 4$. In the continuous case of derivatives such a realisation relation can have the form of an n -th order *differential equation*, expressing the n -th derivative of a state in the terms of other states (where also derivatives up to order $< n$ could occur). So, the question about the existence and purpose of first-order differential equations discussed at the end of Sect. 15.8.3 can be generalized to higher order differential equations.

15.10 Changed Potentialities Due to Interaction

Potentiality can lead to what Aristotle calls ‘the actuality of the potentiality’, e.g., the actual being at position P' , but there may be cases where potentialities are not actualised, but disappear without having their effect. For example, some heavy object can be positioned in such a way that the arrow cannot be at P' , due to its interaction with the object. This section addresses how potentialities can be exchanged between objects by interaction. Some examples are used to show that an interaction can lead to changed potentialities (Sect. 15.10.1). Since changed potentialities can be explained using higher-order potentialities, an interaction can be characterised by the higher-order potentialities it invokes (Sect. 15.10.2). In Sect. 15.10.3 Newton’s laws of mechanics are considered from this perspective. The notion ‘force’ plays the role of a potentiality for the potentiality ‘quantity of motion’ or momentum which itself is a first-order potentiality for distance, so this makes force a second-order potentiality. Within classical mechanics interactions between physical objects are characterised by the forces invoked by the interaction.

15.10.1 Exchange of Potentialities by Interaction

An intensively studied example is one (white) billiard ball A at t' moving to P , while another, equal billiard ball B (red) is positioned at rest in P (see Fig. 15.4). If ball A reaches P at time t , it has the potentiality to be at a next position P' at a next point in time t' . However, what actually occurs is that ball A is still at P at time t' , at rest, and ball B is at P' at time t' .

How can this be explained? A first part of the explanation is that apparently at time t ball B had a potentiality to be at P' ; assuming the presence of this potentiality, the explanation runs as above:

Why is ball B at t' at position P' ?

Ball B is at position P' at t' because

at t it was at position P , and

at t it had the potentiality to be at P' , and

at t nothing in the world excluded it to be at P'

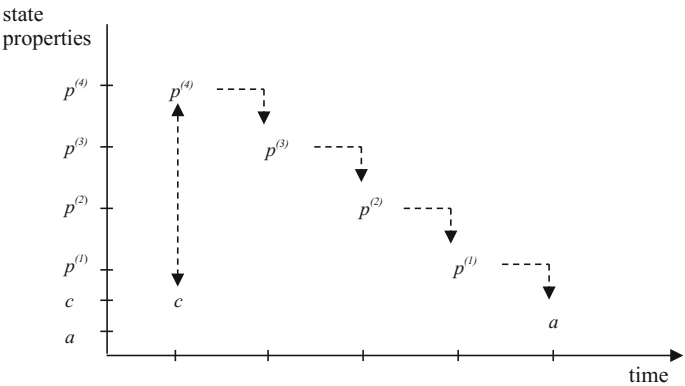
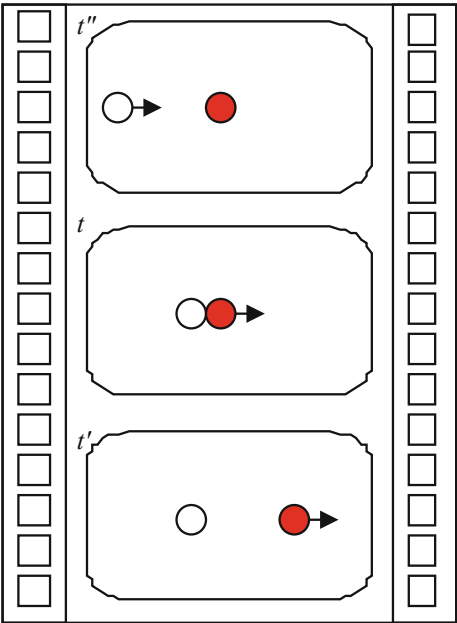


Fig. 15.4 Breaking the chain of higher order potentialities

But how can the presence of the potentiality for ball *B* at time *t* be explained? Apparently potentialities can be transferred from one object to another one: it seems that the potentiality of ball *A* was carried over to a same potentiality of ball *B*, as if a ghost jumps from one body into another body (Fig. 15.5).

A next question is: how can the presence of a potentiality *p* (e.g., of ball *B* at *t*) be explained? Assuming that the billiard ball experiment takes place in isolation of other possible interactions, the only reasonable candidate for the origin of this potentiality *p* is ball *A*, because of its intense interaction (collision) with *B*; a

Fig. 15.5 Transfer of potentiality at time *t* from ball *A* (white) to ball *B* (red)



reasonable explanation is that during this interaction the potentiality p that ball A had before t was transferred to a potentiality p of ball B at t , and thereby ball A lost this potentiality. According to Sect. 15.9 such a change in potentiality p can be explained by a second-order potentiality, say indicated by f :

Why has ball B at t the potentiality to be at position P' ?

At t ball B has the potentiality to be at position P' because

before t ball A had the potentiality to be at position P' , and

at t (due to an interaction) a second-order potentiality f for ball B occurred

Note that a similar pattern can be used to explain that at t the potentiality of ball A was changed to 0, by assuming an opposite second-order potentiality $-f$. So the interaction at t co-occurs with two opposite second-order potentialities on ball A and ball B .

15.10.2 The Role of Higher-Order Potentialities in the Exchange of Potentialities

In Sect. 15.9 higher-order potentialities were introduced to explain changed potentialities. The change of first-order potentialities due to interaction between objects was discussed in Sect. 15.10.1. This suggests how interactions can be characterised using second-order potentialities; see Fig. 15.6. Actually, the interaction as described in Sect. 15.10.1 abstracts from the interaction process itself. It only considers the two states at a time point t just before the interaction and t' just after the interaction and makes up what has changed in the meantime. To be able to explain the first-order potentiality of ball B at t' it has to be assumed that at t second-order potentialities have occurred within both balls. This leads to the assumption that a collision between such (elastic) objects generates second-order potentialities during their contact. How is that possible? Again ghost-like states? Close observation of the physical process of the collision reveals that within the time interval of the interaction at t both balls have some elastic deformation. The idea is that such deformations co-occurs with second-order potentialities: these

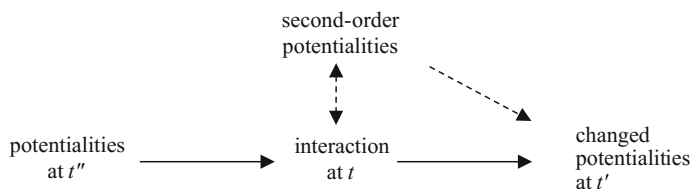


Fig. 15.6 Physical interaction characterised by second-order potentialities; or: second-order potentialities realised by physical interaction

deformations are realisers of the second-order potentialities in the two balls. Such hypothetical second-order potentialities have been called *forces* within Physics.

15.10.3 *Higher-Order Potentialities to Characterise Interaction in Physics*

One of the implications of the conservation law for potentialities in the form of quantity of motion as formulated by Descartes, is that an object in motion and not interacting with other objects remains in (the same quantity of) motion (inertia of motion). This law, already known by Descartes and Galileo, is adopted by Newton (1729) in his *Principia* as the first law. This law states that the absence of impressed forces entails unchanged (quantity of) motion, which suggests that ‘impressed forces’ relate to ‘change of motion’, which makes them second-order potentialities. It is assumed that observations like the colliding balls led Newton (1729) in his *Principia* to reformulate what in principle was already available from, among others, Descartes and Galileo, in the form of what has become known as his third and second law.

Newton’s third law expresses the mutual influence of two objects in interaction by opposite but equal forces (action is reaction). He uses words such as ‘action’, ‘pressing’, ‘drawing’, and ‘equal change of the two motions’. In the case of the colliding balls these are the second-order potentialities indicated above by f and $-f$. In his second law he uses the term ‘impressed motive force’ to express the change of motion. This law expresses that the concept of force used by Newton directly relates to change of motion. Terms like (im)pressing, drawing and action are not further explained. However, for quantity of motion he gives the same definition as Descartes. For an impressed force a definition is given that refers to ‘exerted action’, which itself is not further defined, and to the corresponding change of the object’s state of motion. Furthermore, he shows how this notion applies in the particular case of centripetal (i.e., directed to one point) force.

Newton’s descriptions show that the concept ‘force’ used by him as an addition to the state ontology can be given a definitional relationship to ‘motion generated in a given time’. This ‘motion generated in a given time’ can be considered a second-order potentiality for the first-order potentiality ‘motion’. So, within classical mechanics, after the concepts ‘momentum’ and ‘kinetic energy’ which were added to the state ontology as specific types of (first-order) potentiality, the concept ‘force’ can be considered a third anticipatory state property added to the state ontology, this time as a second-order potentiality. For Newton, initially a force was a discrete event, something that, if repeated, comes in ‘blows’. However, studying the orbits of planets and attempting to explain the circular motion, he had to assume that such blows come all the time with very small time distances between them. To incorporate this and similar phenomena, Newton and also Leibniz developed

mathematical techniques of calculus, such as differentiation and integration. Using these techniques, Newton's second law is formulated as

$$F = \mathbf{dp}/\mathbf{dt} \quad \text{or} \quad F = \mathbf{d}(mv)/\mathbf{dt}$$

For a mass m which is constant over time this is equivalent to

$$F = ma$$

with a the acceleration $\mathbf{dv}/\mathbf{dt}$, which is the second-order derivative $\mathbf{d}^2\mathbf{x}/\mathbf{dt}^2$. In this—most known form—the law was formulated by Euler 65 years after the *Principia* appeared. In 20th century text books such as (Mach 1942) the concept 'moving force' is *defined* in terms of second order potentialities in the following form:

Definition.

Moving force is the product of the mass value of a body with the acceleration induced in that body (Mach 1942), p. 304.

This again shows the second-order potentiality character of a force, defined in terms of acceleration, which is a second-order potentiality for the first-order potentiality velocity.

Analysing the motion of planets around the sun, Newton found out that they can only follow their orbit if a second-order potentiality in them is assumed, in the direction of the sun. So, although there is a very large distance between them, the planets have some mysterious interaction with the sun. Newton calculated (using his calculus under development) in detail that this motive force was proportional to 1 divided by the square of the distance X . For example, for an object in space with mass m at distance X of the earth (with mass M), Newton's law of gravitation for the motive force on the object is as follows (here c is a constant):

$$F = c mM/X^2$$

This can be interpreted as a realisation relation for the second-order potentiality F , taking into account properties of the interaction, such as their distance. It is also possible to rewrite this in the form of a second-order differential equation by substituting $m \mathbf{d}^2X/\mathbf{dt}^2$ for F :

$$\begin{aligned} m \mathbf{d}^2X/\mathbf{dt}^2 &= c mM/X^2 \\ \mathbf{d}^2X/\mathbf{dt}^2 &= c M/X^2 \end{aligned}$$

This shows how the second-order derivative $\mathbf{d}^2X/\mathbf{dt}^2$ for distance X has realiser $c M/X^2$.

But here, the objects being at a large distance, the occurrence of such a second-order potentiality based on some form of interaction is even much more surprising than in the case of the billiard balls, as Nagel states and cites from Newton:

Although it was Newton who propounded the theory of gravitation, he did not regard it as ultimately satisfactory because it involved the notion of ‘action at a distance’ – a notion he regarded as ‘so great an absurdity that I believe no man, who has in philosophical matters a competent faculty in thinking, can ever fall into it’. For he maintained that ‘it is inconceivable that inanimate brute matter should, without the mediation of something else which is not material, operate upon and affect other matter without mutual contact (Newton 1958, pp. 302–303; Nagel 1961 p. 171).

Newton even suggests the option that ‘something else which is not material’, so again some ghost-like entity, could be needed to make such a distant interaction work. This shows that from the perspective of explanation, second-order potentialities realised by a form of interaction between objects still can generate difficult questions, although in this case the physical realisation relation is available.

15.11 Multiple Realisation of Potentialities

A complicating issue for realisation of potentialities is that there may sometimes be a co-occurrence with one other state property and sometimes with another one: *multi-realisability*. Mental state properties can have a large variety of realisers, for example in different animal species. Relating a mental state property in a biconditional manner to all of these mutually distinct (non-equivalent) realisers will lead to a contradiction. If p is equivalent to each of two realisers c_1 and c_2 , then it follows that c_1 is equivalent to c_2 , and thus they always co-occur. In a multiple realisation case where in different states sometimes one, sometimes another realiser co-occurs with p , this is a contradiction.

A solution could be to differentiate the potentiality into a (possibly large) number of distinct variants, thus creating a number of biconditional relationships. However, then the unifying and generic aspect of this potentiality may be lost; (e.g., Kim 1996, pp. 233–236). Therefore, for the case of multi-realisability a broader definition of the notion of realiser is desirable.

In the case of multiple realisers, the relation between potentiality p and its realisers can be described by a supervenience relation (e.g., Kim 1998).

Mental properties supervene over physical properties in that for every mental property M that occurs at some point in time t , there exists some physical property P that also occurs at t , such that always if P occurs at some point in time t' , also M occurs at t' (Kim 1998, p. 9).

This notion can be formulated for potentialities as follows

Multirealisability of a potentiality p by a set of realisers C

At any point in time t ,

p occurs in the state at $t \Leftrightarrow$ there exists a c in C such that c occurs in the state at t

The relation between p and the elements of C is called *multiple realisation relation*. Note that this definition can also be applied to higher-order potentialities. The set of

realisers C can be any specific set of state properties. If this set C can be indicated by a finite number of elements $c_1, \dots, c_n$, then multirealisability can be defined as a bi-implication using a disjunction of the c_i :

$$p \text{ occurs in the state at } t \Leftrightarrow c_1 \vee \dots \vee c_n \text{ holds in the state at } t$$

If this set C just contains one realiser, the standard form of (single) realisation is obtained. Notice that in general the set C , from which multiple realisers come, is a not defined itself by the above definitions. Indeed, in practice this set may be hard to define in a precise manner.

Multiple realisation applied to potentialities, expresses that potentialities are always realised in one way or the other. However, this can happen in a non-systematic, ad hoc manner: for every context a different realiser. This may entail a branching of the potentiality into a multitude of variants, thus loosing the unifying and generic aspect of the potentiality. Sometimes, this situation is avoided by introducing for each context strong, context characterising assumptions excluding all but one of the realisers.

As an example, for a qualitative dynamic modeling approach, multiple realisation can be incorporated easily. Suppose two transitions

$$c_1 \rightarrow d$$

$$c_2 \rightarrow d$$

are given. Then, the set of state properties $C = \{c_1, c_2\}$ can be considered a set of realisers of the potentiality $p(d)$ leading to d .

In Nagel (1961, pp. 186–192), the multiple realization of the notion of force (which can be considered a second-order potentiality; see also Sect. 15.10) is discussed. In line with what was stated above, his analysis asserts that for various different situations specific force-functions, specifying how force relates to other properties of the state are needed. Forces can occur due to state properties involving, for example:

- the presence of an object pushing or pulling
- deformation such as caused by collisions (e.g., billiard balls)
- the presence of objects with electrical charge
- the presence of magnetic objects
- the presence of other masses (gravitation)
- atmospheric pressures.

For each of these circumstances, a different expression in terms of the world state ontology (a force-function) describes a realiser for the force that occurs. Only if for a given situation such a force-function has been identified, something practical can be done using the laws of classical mechanics. In this sense, this case shows a heterogeneous situation, where a force potentiality is described by some heterogeneous disjunctive form with at least, say, up to 5–10 essentially different contexts of the origin of the force. If one aggregated disjunctive realiser would be used this

could lead to weird situations. For example, in such a case to make calculations for the orbit of a satellite around the earth, formulas are to be used not only for gravitation forces but also for collisions, electrical attraction, and so on, all of which do not play a role or are neglectable. For this reason it is more practical to identify the context in which a single realiser is relevant, and only use that realiser in the given context.

If in this heterogeneous situation, in different contexts different force-functions (and hence realisers) are identified, this still allows successful use of the notion of force in applications. This shows an example of a specific approach to multiple realisation, comparable to the notion of local or context-dependent reduction as described by Kim (1996, pp. 211–240).

Notice that also an additive property for this second-order potentiality force holds in the following sense: the combined effect of any number of different contributions to the second-order potentiality can be obtained by adding their values. So, any value w for this second-order potentiality can be obtained as the combined effect from, for example, gravitation, electrical charge, and deformation by collision. For example, considering one dimension where all effects work along the same axis, this can occur in the form of an infinite number of possible sums $w = w_1 + w_2 + w_3$ with the same outcome w , where the terms are the contribution of one of the three effects (e.g., w_1 by gravitation, w_2 by electrical charge, w_3 by collision). This shows that for a given force the complete set of realisers C can be infinite, and also that in reality contexts may occur in complex combinations.

15.12 State-Determined Systems and Potentialities

In this section the main implications of the findings from the previous sections for state-determined systems and temporal-causal networks are briefly summarized. Van Gelder and Port (1995) briefly explain what a dynamical system is in the following manner. A *system* is a set of changing aspects (or state properties) of the world. An (overall) *state* at a given point in time is the way these aspects or state properties are at that time; so an overall state is characterised by the state properties that hold. The set of all possible overall states is the *state space*. A *behaviour* of the system is the change of these state properties over time, or, in other words, a succession or sequence of states within the state space. Such a sequence in the state space can be indexed, for example, by natural numbers (*discrete* case) or real numbers (*continuous* case), and is also called a *trace* or *trajectory*. Given these notions, the notion of *state-determined system*, adopted from Ashby (1960) is taken as the basis to describe what a dynamical system is. In such a system at each point in time the (overall) state fully determines all future states. For more details of state-determined systems and their assumptions, see Chap. 1, Sect. 1.5 and Chap. 2, Sect. 2.9.

Given a particular overall world state that just changed with respect to some of its specific state properties, it is natural to ask for an explanation of why these new state properties occurred. In a not necessarily state-determined system, as a source

for such an explanation, state properties found in the previous state may form a first candidate, with states further back in the past possibly as additional candidates.

For dynamical systems considered as state-determined systems the properties of the previous state are assumed to form the only candidate source, since in such a system the previous state fully determines the next state, so there is no need to look further back in the past, as such past states only have their effect through the previous state. Thus, a main question becomes how to determine on the basis of some of the specific state properties in the given overall state, the specific state properties in the new state. More specifically, can particular state properties (or combinations of state properties) occurring in a given state be identified that in some way or the other indicate the (changed) state properties occurring in a subsequent state. By having these particular properties the overall state anticipates on the next state: as briefly discussed above, these properties have been considered historically as anticipatory state properties or potentialities. Due to them, anticipation to change is somehow encoded in a state. The existence of such properties is a crucial factor for the validity of the assumptions underlying the Dynamical Systems Theory.

Recall from Chap. 2, Sect. 2.9 that a continuous system with states X_i is a state-determined system if and only if it can be described by a set of first-order differential equations

$$\mathbf{d}X_i(t)/\mathbf{d}t = f_i(X_1(t), \dots, X_k(t))$$

Through this the analysis above shows that any state-determined system is basically a specification of realisation relations for all potentialities $\mathbf{d}X_i(t)/\mathbf{d}t$ in terms of the states $X_i(t)$. If these states themselves are genuine state properties, then this solves the ghost-like causation problem. For example, what is the cause that $X_i(t)$ has a changed value at $t + \Delta t$? The answer now is that although the potentiality $\mathbf{d}X_i(t)/\mathbf{d}t$ indicates this changed value in a generic manner by

$$X_i(t + \Delta t) = X_i(t) + \mathbf{d}X_i(t)/\mathbf{d}t \Delta t$$

by the realisation relation $\mathbf{d}X_i(t)/\mathbf{d}t = f_i(X_1(t), \dots, X_k(t))$ this can be replaced by

$$X_i(t + \Delta t) = X_i(t) + f_i(X_1(t), \dots, X_k(t)) \Delta t$$

so then it is in fact $f_i(X_1(t), \dots, X_k(t))$, which is a state property for the state at time t , that is causing the change of $X_i(t)$. This means that the causality driving the dynamics in the system is from $f_i(X_1(t), \dots, X_k(t))$ and $X_i(t)$ to $X_i(t + \Delta t)$. A similar analysis applies to the temporal-causal format introduced in Chap. 2:

$$\mathbf{d}X_i(t)/\mathbf{d}t = \eta_i[c_i(\omega_{1,i}X_1(t), \dots, \omega_{k,i}X_k(t)) - X_i(t)]$$

Based on this, it is, more specifically, the state property $\eta_i [c_i(\omega_{1,i}X_1(t), \dots, \omega_{k,i}X_k(t)) - X_i(t)]$ which indicates the change of $X_i(t)$. This ‘indication of change’

corresponds to the aggregated impact by the assumed causal relations from $X_j(t)$ to $X_i(t)$ with strengths $\omega_{j,i}$ and combination function $c_i(\dots)$.

Such an interpretation based on causal relationships is indeed how in Chap. 2 it was described how state-determined systems can be modeled both by a conceptual and a numerical representation. This provides a solid basis for any state-determined system as long as either all states concern genuine state properties or they have realisers, i.e., they are equivalent to genuine state properties.

It may still be the case that in a state-determined system one or more of the variables X_i is not a genuine state property, nor can be related to a realiser. A simple example of this is the following system describing a falling object:

$$\begin{aligned} \mathbf{d}X_1(t)/\mathbf{d}t &= X_2(t) & \text{so } f_1(X_1(t), X_2(t)) &= X_2 \\ \mathbf{d}X_2(t)/\mathbf{d}t &= 10 & \text{so } f_2(X_1(t), X_2(t)) &= 10 \end{aligned}$$

Here $X_1(t)$ is the vertical distance at time t and $X_2(t)$ can be interpreted as the velocity at time t , which is linearly increasing over time, due to gravitation. The latter property can be considered a potentiality for distance and not a genuine state property, as it has no realizer. But it still has another type of relation to the world describing how it changes over time, by the second equation, which expresses that the potentiality $\mathbf{d}X_2/\mathbf{d}t$ for this first-order potentiality X_2 has a realizer, namely the constant 10. This potentiality $\mathbf{d}X_2/\mathbf{d}t$ for potentiality X_2 for X_1 is a second-order potentiality for X_1 . The differential equations can also be rewritten into a second-order differential equation for the second-order derivative $\mathbf{d}^2X_1(t)/\mathbf{d}t^2$ of $X_1(t)$:

$$\mathbf{d}^2X_1(t)/\mathbf{d}t^2 = 10$$

These higher-order potentialities and their role in Physics have been discussed in more detail in Sects. 15.9 and 15.10.

15.13 Discussion

In this chapter the focus was on a philosophical and historical reflection on dynamics, in relation to the basic assumptions underlying state-determined systems (Ashby 1960). The text mainly follows (Treur 2016); part of this content was adopted from Treur (2005). The notion of a state-determined system is central for the Network-Oriented Modeling approach based on temporal-causal networks that forms the core approach of this book. The basic assumption of a state-determined system is that each state of the system fully determines the system's next states, or formulated from a different angle, each state of such a system is fully determined by the system's previous state. Some of the state properties of this previous state in a sense anticipate on the changed state properties in the current state: the changed

properties can be predicted from them. This idea plays an important role in the analysis of dynamics in history; such anticipatory state properties are often called potentialities.

In this chapter this perspective on dynamics was discussed and illustrated for different contexts varying from moving physical objects (e.g., Zeno, Aristotle, Newton) to animal or human action (e.g., Aristotle) and computer systems. It was shown how in history the perspective based on potentialities has led to a number of often used concepts within classical mechanics, a branch of Physics frequently used by engineers today: momentum, energy, and force. Also it was discussed how within mathematics, more specifically in calculus, a number of concepts have been developed to formalise notions of potentiality: in particular derivatives (of different orders) of a function and Taylor approximations. Furthermore, it was shown how causal relations and transition systems, a currently (within Artificial Intelligence, Computer Science and related areas) popular format for specification of dynamic systems can be interpreted from the perspective of potentialities.

Within Cognitive Science one of the problems identified is the problem of realism, i.e., how do internal mental states relate to the real world in a natural manner. This issue also applies to potentialities as assumed state properties. If there is no relation to the physical world they seem to have a ghost-like character, which is not desirable, especially when it would make their occurrence unpredictable. The question in how far such assumed state properties are genuine or 'real' state properties was shown to be a hard question that is not simple to answer in general, even not in Physics. Nevertheless, the fruitfulness of having such added state properties is uncontroversial; for example we all trust artefacts in our environment that were constructed based on Physics and Mathematics using such state properties.

More specifically, potentialities as postulated state properties may have relationships to other state properties of the state in which they occur; they can be said to be realised by these other state properties. But such *realisers* do not always exist, which may leave potentialities with a ghost-like character. However, there is a second way in which potentialities can relate to other state properties, although not in the same state. They often can be related, in a temporal manner, to state properties in other (past and future) states; this corresponds to Kim (1996)'s notion of (temporal) relational specification of mental state properties. These realisation relationships within one state and temporal relationships between states can be exploited to obtain predictability of potentialities. So, in cases that these state properties are felt as ghost-like, at least they are predictable ghosts. This predictability makes them still useful in many types of applications.

One often used way in which it can be specified how potentialities indicating the change of a state relate to other state properties in the same state is by first-order differential equations. A first-order differential equation for a state X is an expression that postulates that at each point in time t the value of the derivative of X at t is equal to a mathematical expression (function) in terms of values of the other states at t . When these states are genuine, this addresses the realism problem for dynamical systems, namely by specifying them by first-order differential equations.

Or, conversely, the abundantly used way of modeling by first-order differential equations can be (re)interpreted as a way of creating dynamics without allowing ghost-like states. Recall from Chap. 2 that a smooth dynamical system is a state-determined system if and only if it can be described by a set of first-order differential equations if and only if it can be described by a temporal-causal network. So, the above analysis applies to any smooth state-determined system and any temporal-causal network, with genuine states: their dynamics can be described without having to rely on ghost-like properties.

References

- Aristotle, *Physica* (translated by R.P. Hardie and R.K. Gaye) (350 BC a)
 Aristotle, *De Motu Animalium* On the motion of animals (translated by A.S.L. Farquharson) (350 BC b)
 Aristotle, *Nicomachean ethics* (translated by W.D. Ross) (350 BC c)
 R. Ashby, *Design for a Brain* (Chapman & Hall, London, 2nd edn, 1960)
 T. Bosse, J. Treur, Higher-order potentialities and their reducers: a philosophical foundation unifying dynamic modelling methods, in *Proceedings of the Twentieth International Joint Conference on Artificial Intelligence, IJCAI'07*, ed. by M.M. Veloso (AAAI Press, 2007), pp. 262–267
 T. Bosse, J. Treur, A philosophical foundation for unification of dynamic modelling methods based on higher-order potentialities and their reducers. *Adv. Complex Syst. J.* **11**, 831–860 (2008)
 P.T. Burns, The complete history of the discovery of cinematography (2000). <http://precinemahistory.net/>
 R. Descartes, *Principles of philosophy* (translated by M.S. Mahoney) (1644)
 R. Descartes, The world or treatise on light, in *Descartes: The World and Other Writings*, ed. by S. Gaukroger (Cambridge University Press, 1998) (translated by M.S. Mahoney)
 H.G. Frankfurt, Freedom of the will and the concept of a person. *J. Philos.* **68**, 5–20 (1971)
 T.J. Van Gelder, R.F. Port, *It's About Time: An Overview of the Dynamical Approach to Cognition* (Port and van Gelder, 1995), pp. 1–43
 T.L. Heath, *A History of Greek Mathematics*, vol. 1 (Oxford, 1931)
 J. Kim, *Philosophy of Mind* (Westview press, 1996)
 J. Kim, *Mind in a Physical world: An Essay on the Mind-Body Problem and Mental Causation* (MIT Press, Cambridge, Mass, 1998)
 G.W. Leibniz, A memorable error of Descartes (1686a)
 G.W. Leibniz, Discourse on metaphysics (1686b)
 G.W. Leibniz, Les surprises du Phoronomus. *Les Études Philosophiques* (April–June), 171–86 (1989)
 G.W. Leibniz, Phoronomus seu De potentia et legibus naturæ. *Dialogus II. Physis* **28**, 797–885 (1991)
 E. Mach, *Science of mechanics*. Peru, IL: Open Court (1942)
 E. Nagel, *The Structure of Science* (Routledge and Kegan Paul; Harcourt, Brace and World, London, 1961)
 I. Newton, *The Mathematical Principles of Natural Philosophy; Newton's Principles of Natural Philosophy* (Dawsons of Pall Mall, 1729)
 I. Newton, *Isaac Newton's Papers and Letters on Natural Philosophy*, ed. by B. Cohen (Cambridge, Mass, 1958)
 M. Nussbaum (ed.), *Aristotle's De Motu Animalium* (Princeton University Press, Princeton, 1978)

- R.F. Port, T. van Gelder (eds.), *Mind as Motion: Explorations in the Dynamics of Cognition* (MIT Press, Cambridge, Mass, 1995)
- J. Sachs, Aristotle: motion and its place in nature. The internet encyclopedia of philosophy (2001). <http://www.utm.edu/research/iep/a/aris-not.htm>
- J. Tennenbaum, Leibniz and dynamics: motion is not simple. Readings from the American Almanac. The new federalist (2002). http://members.tripod.com/~american_almanac/dynamics.htm
- J. Treur, States of change: explaining dynamics by anticipatory state properties. *Philos. Psychol. J.* **18**, 441–471 (2005)
- J. Treur, *What Is It that Drives Dynamics: We Don't Believe in Ghosts, Do We? Transactions on Computational Collective Intelligence* **23**, 212–250 (2016)

Chapter 16

Making Smart Applications Smarter

Societal Applicability of Computational Models

Abstract This chapter briefly outlines how dynamic computational models, and in particular temporal-causal network models, can contribute to smarter applications. The scientific area that addresses Ambient Intelligence (also called Pervasive Computing) applications is discussed in which both sensor data and knowledge from the human-directed sciences such as health sciences, neurosciences, and psychological and social sciences are incorporated. This knowledge enables the environment to perform more in-depth, human-like analyses of the functioning of observed humans, and to come up with better informed actions. It is discussed which ingredients are important to realise this view, and how frameworks can be developed to combine them to obtain the intended type of systems: coupled reflective human-environment systems. Such systems include computational models by which they are able to model and simulate (parts of) their own behavior. Finally, further perspectives are discussed for Ambient Intelligence applications based on these coupled reflective systems.

16.1 Introduction

For societal applicability of computational models and in particular those designed by Network-Oriented Modeling as discussed in this book, the area of Ambient Intelligence (also called Pervasive Computing) is a relevant area. Within this area smart applications are developed in which both sensor data and knowledge from the human-directed sciences such as health sciences, neuroscience, and psychological and social sciences are incorporated. Often the smartness is mainly based on the availability of sensor data, and the knowledge used is limited, and often simplified (or compiled) to direct functional associations or (heuristic) rules of thumb. In other words, often only shallow knowledge models are used and not deep models that involve the underlying causal relations of the human-related domain that is

concerned; for this distinction, see, e.g. (Chandrasekaran and Mittal 1982; Dhar and Pople 1987; Davis 1983).

Ambient Intelligence is important for society in that it provides possibilities to contribute to more personal care; e.g., (Aarts et al. 2001, 2003; Riva et al. 2005; Sadri 2011; Acampora et al. 2013). Acquisition of sensor information about humans and their functioning is an important factor, but without adequate knowledge for analysis of this information, the scope of such applications is limited. However, devices in the environment possessing such knowledge can show a more human-like understanding and are more socially aware, and base personal care on this understanding and awareness. For example, this may concern elderly people, patients depending on regular medicine usage, surveillance, penitentiary care, psychotherapeutical/selfhelp communities, but also, for example, humans in highly demanding tasks such as warfare officers, air traffic controllers, crisis and disaster managers, and humans in space missions; e.g., (Green 2005; Itti and Koch 2001).

Within human-directed scientific areas, such as cognitive science, psychology, neuroscience and health sciences, causal knowledge has been and is being developed for a variety of aspects of human functioning. If such causal knowledge of human processes is represented as causal models in a formal and computational format, and incorporated in the human environment in devices that monitor the physical and mental state of the human, then such devices are able to perform a more in-depth analysis of the human's functioning. This can result in an environment that may more effectively affect the state of humans by undertaking in a knowledgeable manner actions that improve their wellbeing and performance. For example, the workspaces of naval officers may include systems that, among others, track their eye movements and characteristics of incoming stimuli (e.g., airplanes on a radar screen), and use this information in a computational model that is able to estimate where their attention is focussed at. When it turns out that an officer neglects parts of a radar screen, such a system can either indicate this to the person, or arrange on the background that another person or computer system takes care of this neglected part. In applications of this type, an ambience is created that has a better understanding of humans, based on computationally formalised knowledge from the human-directed disciplines.

In this chapter, in Sect. 16.2 it is discussed how multiple disciplines have to play a role as ingredients, and in Sect. 16.3 it is discussed how these ingredients are integrated in order to obtain an ambient intelligent application with human-like understanding which is socially aware of its social environment. Section 16.4 zooms in at the underlying coupled reflective system architecture to describe such systems. In Sect. 16.5 it is addressed in more detail how different types of models can be integrated to fill such a coupled reflective architecture. Finally, Sect. 16.6 is a discussion.

16.2 Multidisciplinarity: The Ingredients

The area as sketched is essentially multidisciplinary. It combines aspects of Ambient Intelligence with knowledge from human-directed disciplines such as psychology, social science, neuroscience and biomedical sciences. Further development will depend on cooperation between researchers from these disciplines or working on cross connections of Ambient Intelligence with the human-directed disciplines. The focus is on the use of knowledge from these disciplines in Ambient Intelligence applications, in order to take care in a more sophisticated manner of humans in their daily living in medical, psychological and social respects. For example, modelers in the psychological, neurological, social or biomedical disciplines interested in Ambient Intelligence as a high-potential application area for their models, can get inspiration for problem areas to be addressed for further developments in their disciplines. From the other side, researchers in Computer Science, and Artificial and Ambient Intelligence may become more aware of the possibilities to incorporate more substantial knowledge from the psychological, neurological, social and biomedical disciplines in Ambient Intelligence architectures and applications, and may offer problem specifications that can be addressed by the human-directed sciences.

In more detail, content from the domain of human-directed sciences, among others, can be taken from areas such as medical physiology, health sciences, neuroscience, cognitive psychology, clinical psychology, psychopathology, sociology, criminology, and exercise and sport sciences. From the domain of Artificial Intelligence, useful contributions can be found in areas such as knowledge and task modeling, and cognitive and social modeling and simulation. Finally, from the Computer Science domain, relevant areas are distributed systems, sensor systems, human-centred software engineering, user modeling, and human-computer interaction.

16.3 Combining the Ingredients

One of the challenges is to provide frameworks that cover the class of Ambient Intelligence applications being socially aware and showing human-like understanding and supporting behaviour. Here human-like understanding is defined as understanding in the sense of being able to analyse and estimate what is going on in the human's mind (a form of mindreading) and in his or her body (a form of bodyreading). Input for these processes are observed information about the human's state over time, and dynamic models for the human's physical and mental processes. For the mental side such a dynamic model is sometimes called a Theory of Mind (e.g., Baron-Cohen 1995; Dennett 1987; Gärdenfors 2003; Goldman 2006)

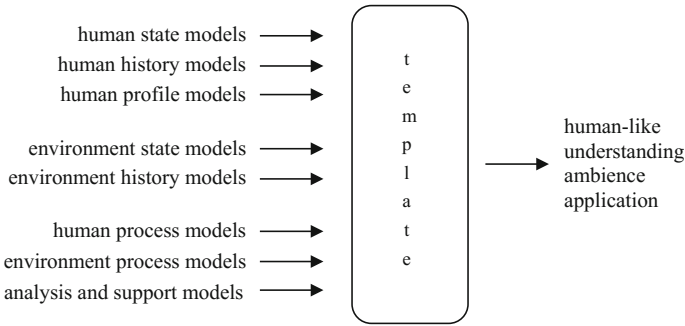


Fig. 16.1 Framework to combine the ingredients

and may cover, for example, emotion, attention, intention, and belief. Similarly for the human's physical processes, such a model relates, for example, to skin conditions, heart rates, and levels of blood sugar, insulin, adrenalin, testosterone, serotonin, and specific medicines taken. Note that different types of models are needed: physiological, neurological, cognitive, emotional, social, as well as models of the physical and artificial environment.

A framework can be used as a template for the specific class of Ambient Intelligence applications as described. The structure of such an ambient software and hardware design can be described at a conceptual design level and can be given generic facilities built into represent the following (see also Fig. 16.1):

- human state and history models
- environment state and history models
- profiles and characteristics models of humans
- ontologies and knowledge from biomedical, neurological, psychological and/or social disciplines
- dynamic process models about human functioning
- dynamic environment process models
- methods for analysis on the basis of such models and for support by interventions

Examples of such analysis methods are voice and skin analysis with respect to emotional states, gesture analysis, heart rate analysis. The template can include slots where the application-specific content can be filled to get an executable design for a working system. This specific content together with the generic methods to operate on it, provides an overall system, based on a tight cooperation between a human and an ambient system which is socially aware and has human-like understanding of its social environment and reacts from this understanding and awareness in a knowledgeable manner.

16.4 Coupled Reflective Systems

Ambient Intelligence applications in general can be viewed as coupled reflective (human-environment) systems, where ‘coupled’ means mutually interacting. For the specific type of applications considered here, however, the coupling takes two different forms; see also Fig. 16.2.

- On the one hand the coupling takes place as interaction between human and environment, as in any Ambient Intelligence application:
 - the environment gets information generated by the human as input, and
 - the human gets information generated by the environment as input.
- In addition, coupling at a more deep, reflective level takes place due to the fact that
 - the environment has and maintains knowledge about the functioning of the human, the environment and their interaction, and
 - the human has and maintains knowledge about functioning of him or herself, the environment, and their interaction

So, in such a more specific human-environment system, being coupled does not only mean that the human and its environment interact, but also that they have knowledge, understanding and awareness of each other, themselves and their interaction. This entails two types of awareness:

- *Human awareness*: awareness by the human about the human and environmental processes and their interaction
- *Technological awareness*: awareness by the environment about the human and environmental processes and their interaction

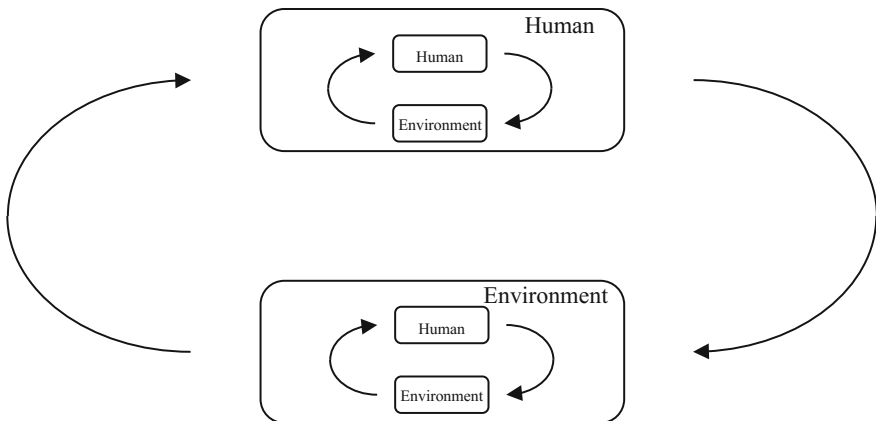


Fig. 16.2 Coupled reflective system architecture

By (human and technological) learning, adaptation and development processes for both the human and the environment, these types of awareness can also grow over time.

Such a coupled reflective system can have a positive impact at different aggregation levels, from individual via an organisation within society to the society as a whole:

- *Individual level*
 - more effective functioning
 - stimulating healthy functioning and preventing health problems to occur
 - support of learning and development
- *Organisation level*
 - efficient functioning organisation by wellfunctioning members
 - learning and adaptation of the organisation
- *Society level*
 - limiting costs for illness and inability to work
 - efficient management of environment

In Sect. 16.5 it will be discussed how this global architecture of a coupled reflective system can be filled with integrated computational models in order to obtain an ambient intelligence application.

16.5 Integrative Modeling

A coupled reflective architecture is a blueprint of a system at a very global level. In this section different types of computational models are distinguished that can be integrated within this coupled reflective architecture to get an actual design for such a system.

Domain and human models are used to describe phenomena in the natural and human world such as for example:

- the physical and biological environment
- physiological and cognitive processes within humans
- behavioural processes of a human in interaction with his or her environment
- social and organisational processes based on interaction between humans

Models for software applications can be used to describe artificially created (engineered) processes such as, for example:

- smartphone apps
- intelligent cars
- smart homes

- search bots and recommender systems
- robots

To model human functioning, models can be designed, for example, for humans with behaviour motivated by desires, and humans generating emotions and feelings. Moreover, domain models can be integrated within ambient application models in order to give these ambient applications some understanding of the part of reality they are dealing with. Such understanding enables them to perform actions in a more informed, knowledgeable manner, and to show more human-like behaviour in interaction with humans. For more a extensive exposition, see (Bosse et al. 2011a, b, 2012, 2013).

The required integration takes place by embedding domain models in certain ways within application models. By incorporating domain models within an application model, the application gets an understanding of the processes of its surrounding environment, which is a solid basis for knowledgeable intelligent behaviour. Four different ways to integrate domain models within application models are considered here. A most simple way is to use a domain model that specifically models human behaviour in the following manner:

- *domain model referring to or simulating the human process*
A domain model describes human processes and behaviour and as such refers to them. The domain model can also be used to simulate human behaviour.

Such a model can be used in interaction with other models, in particular with models for smart applications to obtain a test environment for simulations. For this last type of models, domain models can be integrated in three different ways, in order to obtain the following (sub)models; see Fig. 16.3. Here the solid arrows indicate information exchange between processes (data flow) and the dotted arrows the integration process.

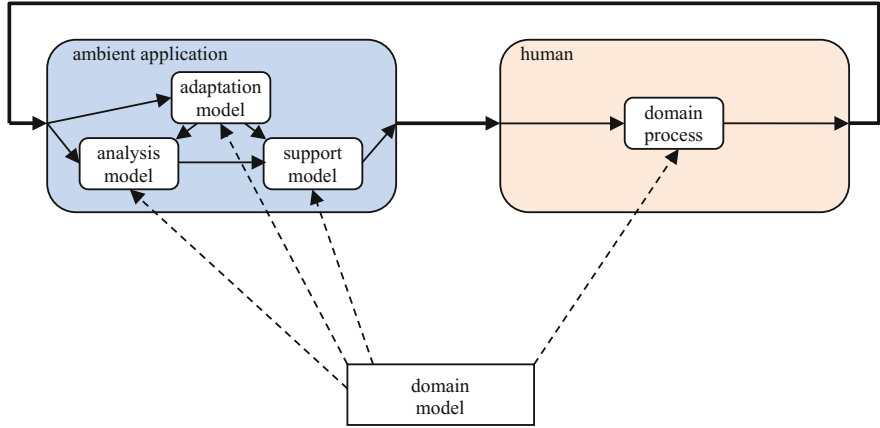


Fig. 16.3 Integrative model

- *analysis model*
To perform analysis of the human's states and processes by simulation (of the human-environment interaction) and reasoning based on observations (possibly using specific sensors) and the domain model.
- *support model*
To generate support for the human by (what-if) simulation (of the human-environment interaction) and reasoning based on the domain model.
- *adaptation model*
To tune parameters in the domain model better to the specific characteristics of the humans by reasoning based on the domain model.

Some more specific examples of today's societal challenges, to which coupled reflective human-environment systems can contribute, are elderly care, health management, crime and security.

16.6 Discussion

Parts of the content of this chapter are based on: (Treur 2008). The scientific area that addresses Ambient Intelligence applications in which knowledge from the human-directed sciences is incorporated, has a high potential to provide nontrivial Ambient Intelligence applications based on human-like understanding. Such understanding can result in better informed actions and will feel more natural for humans. Important additional ingredients to realise this view are provided by areas in Computer Science, Artificial Intelligence and Cognitive Science; among others: knowledge and task modeling, user modeling, and cognitive and social modeling. Furthermore integrative frameworks can be developed to combine the ingredients. The resulting human-environment systems are coupled not only by their mutual interaction, but also in a reflective manner in the sense that both the human and the ambient system have and/or develop a model of the interactive processes of the human and the environment.

These coupled reflective human-environment systems are an interesting type of systems to be studied scientifically, and provide a solid foundation for human-like, socially aware Ambient Intelligence applications. It has been pointed out how such applications can be designed in a principled manner based on integrative modeling of an ambient applications involving a computational domain model embedded in an analysis model, a support model and an adaptation model as components; see also (Bosse et al. 2011a, b, 2012, 2013). These elements facilitate development of human-like Ambient Intelligence applications with significant benefits for individuals, organisations, and the society as a whole. Dynamic computational models such as temporal-causal network models play a crucial role in this development.

References

- E. Aarts, R. Harwig, M. Schuurmans, Ambient Intelligence, in *The Invisible Future*, ed. by P. Denning (McGraw Hill, New York, 2001), pp. 235–250
- E. Aarts, R. Collier, E. van Loenen, R. de Ruyter (eds.), Ambient intelligence, in *Proceedings of the First European Symposium, EUSAI 2003. Lecture Notes in Computer Science*, vol. 2875 (Springer, 2003), pp. 432
- G. Acampora, D.J. Cook, P. Rashidi, A.V. Vasilakos, A survey on ambient intelligence in healthcare. *Proc. IEEE* **101**(12), 2470–2494 (2013)
- S. Baron-Cohen, *Mindblindness* (MIT Press, 1995)
- T. Bosse, M. Hoogendoorn, M.C.A. Klein, R.M. van Lambalgen, P.P. van Maanen, J. Treur, Incorporating human aspects in ambient intelligence and smart environments, in *Handbook of Research on Ambient Intelligence and Smart Environments: Trends and Perspectives. IGI Global*, ed. by N.Y. Chong, F. Mastrogiovanni, pp. 128–164 (2011a)
- T. Bosse, M. Hoogendoorn, M.C.A. Klein, J. Treur, An ambient agent model for monitoring and analysing dynamics of complex human behaviour. *J. Ambient Intell. Smart Environ.* **3**, 283–303 (2011b)
- T. Bosse, F. Both, C. Gerritsen, M. Hoogendoorn, J. Treur, Methods for model-based reasoning within agent-based ambient intelligence applications. *Knowl.-Based Syst. J.* **27**, 190–210 (2012)
- T. Bosse, F. Both, R. Duell, M. Hoogendoorn, R. van Lambalgen, M.C.A. Klein, A. van der Mee, R. Oorburg, A. Sharpanskykh, J. Treur, M. de Vos, An ambient agent system assisting humans in complex tasks by analysis of a human's state and performance. *Int. J. Intell. Inf. Database Syst.* **7**, 3–33 (2013)
- B. Chandrasekaran, S. Mittal, Deep versus compiled knowledge approaches to diagnostic problem-solving. *Proc. AAAI-82*, 349–354 (1982)
- R. Davis, Reasoning from first principles in electronic troubleshooting. *Int. J. Man-Mach. Stud.* **19**, 403–423 (1983)
- D.C. Dennett, *The Intentional Stance* (MIT Press, Cambridge Mass, 1987)
- V. Dhar, H.E. Pople, Rule-based versus structure-based models for explaining and generating expert behaviour. *Commun. ACM* **30**, 542–555 (1987)
- P. Gärdenfors, *How Homo Became Sapiens: On The Evolution Of Thinking* (Oxford University Press, 2003)
- A.I. Goldman, *Simulating Minds: The Philosophy, Psychology and Neuroscience of Mind Reading* (Oxford University Press, Oxford, 2006)
- D.J. Green, realtime compliance management using a wireless realtime pillbottle—a report on the pilot study of SIMPILL, in *Proceedings of the International Conference for eHealth, Telemedicine and Health, Med-e-Tel'05* (Luxemburg, 2005)
- L. Itti, C. Koch, Computational modeling of visual attention. *Nat. Rev. Neurosci.* **2**(3), 194–203 (2001)
- G. Riva, F. Vatalaro, F. Davide, M. Alcañiz (eds), *Ambient Intelligence* (IOS Press, 2001)
- F. Sadri, Ambient intelligence: a survey. *ACM Comput. Surv. (CSUR) Surv.* **43**(4) (Article No. 36) (2011)
- J. Treur, On human aspects in ambient intelligence, in *Proceedings of the First International Workshop on Human Aspects in Ambient Intelligence*, ed. by M. Muehlhauser, A. Ferscha, E. Aitenbichler, Constructing Ambient Intelligence: Aml-07 Workshops Proceedings. Communications in Computer and Information Science (CCIS), vol. 11 (Springer, 2008) pp. 262–267

Chapter 17

Multidisciplinary Education

Computational Modeling as the Core of a Multidisciplinary Curriculum

Abstract This chapter discusses the design of a curriculum with main focus on human-oriented scientific knowledge and how this can be exploited to develop support for humans by means of advanced smart devices in the daily environment. The aim for this curriculum was to offer a study path for those students with exact talents but with an interest mainly in human processes and society. The curriculum was designed from a problem-oriented perspective in relation to societal problem areas. From human-oriented disciplines scientific knowledge for human processes in such problem areas was obtained. Computational modeling for such human processes plays a central role as an integrating factor in the curriculum. Elements from Ambient Intelligence, Artificial Intelligence, and Informatics are included for design of smart support systems.

17.1 Introduction

As discussed in Chap. 16, computational modeling is an important ingredient to make smart applications smarter than they often are currently. To develop such smart applications requires knowledge and skills from a number of domains and their integration, for example, on the one hand health sciences, psychology and neurosciences and social sciences, and on the other hand more engineering-directed sciences such as Computer Science and AI. These types of expertise could be brought together in multidisciplinary teams where each member represents a single discipline. However, communication in such teams often is difficult. Therefore it makes sense to also educate persons in a multidisciplinary manner, so that they know different disciplines and know how to integrate them, and how to communicate with experts from such disciplines. This chapter discusses how a curriculum for such multidisciplinary education can be designed. In particular, such a curriculum has to bring together subjects from human-directed sciences and more technical engineering-directed sciences.

In general a challenging issue in academic education is how to interest candidate students in the further development of their talents in more technical, engineering

sciences. The numbers of students choosing such a exact study are usually very small compared to the numbers welcomed for studies, for example, in Health Sciences, Psychology or Social Sciences, and this trend is even still much stronger for female students in particular. Among this large group of students a substantial subgroup is made up of students whose talents are not in the first place in exact, more formal and technical subjects. However, there exists also a substantial subgroup consisting of students with good talents for more exact scientific work, but whose interest is simply not in these topics, but rather in human processes and society. Currently, individuals from the latter subgroup do not further develop their talents for exact sciences, which is a pity both for them and for society in general. Actions taken in the past to create more advertising for exact sciences have not brought much change to this situation. The curriculum design presented in this chapter took as a point of departure the hypothesis that the available curricula in exact sciences are not satisfactory for students whose main interest is in humans and society. Therefore the question was addressed as to how an academic curriculum can be designed which is attractive for students with both exact talents and an intrinsic interest in humans and their functioning in society.

The aim was to develop a 5-year curriculum [Bachelor (3 years) and Master (2 years)] in which the main focus is on human-directed scientific knowledge (from health, psychological and social sciences, indicated here as human sciences) and on how—using elements from exact sciences—this can be exploited to develop scientifically-justified support for humans by means of advanced smart devices (such as smartphones) in the day-to-day environment. The idea was that this would provide a study path which is attractive for those students with an interest in human functioning and society who also have exact talents, and for female students in particular.

The curriculum was meant to provide a new, broad, multidisciplinary study focusing on human functioning in physical, mental and social respects. Human wellbeing and functioning depends on many factors in the environment. This environment can contribute positively (e.g. a workplace avoiding RSI, a nice living room), but it can also have negative effects (e.g. too-high work demands, disturbances during sleep). Insight into interaction between humans and their environment makes it possible to stimulate the positive aspects and limit the negative ones. In this curriculum, insights into human functioning are acquired. Moreover, it is learnt how these insights can be applied to various practical problems, how such problems can be analysed, and how solutions can be designed by making use of supporting devices so that a more understanding environment is created. This may concern, for example, microphones that can determine whether fear or aggression is present during an encounter, a wrist belt for elderly people that can detect medical early warning signs, or a car that notices when a driver is drunk or may fall asleep. After this study students may be employed by the R&D departments of companies that develop such modern technological devices and focus on the knowledge and models applied in such devices.

The choice was made to design the curriculum from a problem-oriented perspective. Examples of societal problem areas chosen include supporting patients with chronic diseases (e.g. diabetes) or mental problems (e.g. mood disorders), care for elderly persons in their living environment, support for persons in demanding circumstances (e.g. sportspersons, air traffic controllers). Subjects in human-directed disciplines that provide scientific knowledge for human functioning in such areas, such as those offered by other faculties in biomedical, psychological and social sciences, were identified. To create a bridge from these informal, nontechnical bodies of knowledge to the exact domain, specific subjects were developed addressing the computational modeling of such human processes, thereby using formal, computational modeling techniques from Computational Science, Artificial Intelligence and Informatics. Moreover, from the areas of Ambient Intelligence, Artificial Intelligence, and Informatics, subjects were developed that demonstrated how to integrate computational models, based on scientific knowledge of human processes, with sensor systems and intervention methods to obtain support in a knowledgeable, human-aware manner, for example through a smartphone.

In a few years, a problem-oriented, multidisciplinary curriculum (3-year Bachelor and 2-year Master) was successfully developed along these lines, integrating human sciences with exact sciences, which seemed to be unique. For practical reasons the new curriculum was developed to replace an existing curriculum in Artificial Intelligence. The results of this so far are that the newly-developed curriculum has to date attracted substantially more (up to a factor 2) students than the original Artificial Intelligence curriculum. Moreover, an investigation was carried out in a large number of high schools as to how many students (in their penultimate year) would be interested in choosing for such a curriculum in the future. The outcome of this was that 66 % had some or much interest in choosing such a curriculum (boys 60 %, girls 75 %). It was shown to attract much more interest from students than the more traditional curriculum it was replacing.

In this chapter, Sect. 17.2 describes the overall structure of the designed curriculum. In the subsequent Sects. 17.3–17.5 the four main streams in the curriculum are discussed in more detail. Section 17.6 addresses evaluation and discussion.

17.2 Overall Structure of the Curriculum

The aim to design a problem-oriented, multidisciplinary curriculum does not only entail that ingredients from different disciplines are to be incorporated along with different social problem areas, but also that these ingredients have to be integrated in a certain way made to contribute to these societal problem areas. It will be clear that it does not suffice just to include different subjects from a number of disciplines in the curriculum with the idea that students will integrate and use these subjects by themselves in the application areas at hand. To obtain an effective curriculum is

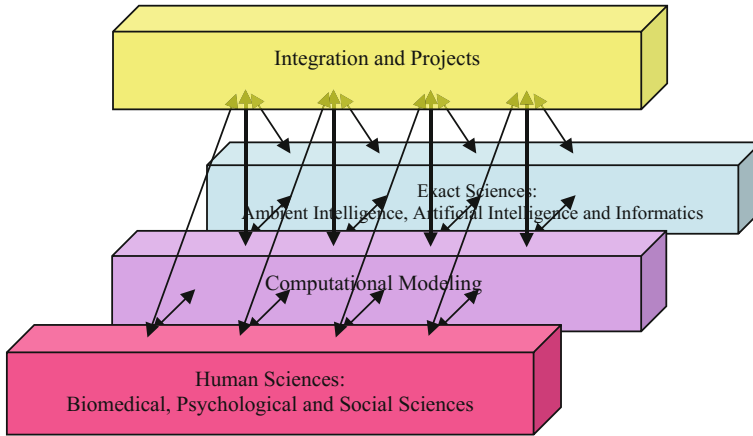


Fig. 17.1 The four main streams in the curriculum and their main interactions

quite challenging, and requires much attention to integration, analysis and application with respect to the problem areas.

The four main streams and their interactions are depicted in Fig. 17.1. The following gives an impression of their approximate relative size in the design of the curriculum: Human Sciences stream (30–35 %), Exact Sciences stream (25 %), Computational Modeling stream (20–25 %), and Integration and Projects stream (20 %).

The Human Sciences stream and the Exact Sciences stream cover relevant topics from biomedical, psychological and social sciences, and from ambient intelligence, artificial intelligence, and informatics, respectively. As integration is a crucial element in all this, and the different scientific disciplines used as ingredients differ enormously, first of all a stream on *Computational Modeling* was included which focuses on the integration of human sciences and exact sciences. This stream serves as a strong integration factor as it is here that students learn to take (informally described) topics from the *Human Sciences* stream on the one hand and methods and techniques from the *Exact Sciences* stream on the other and glue them together in a formalised computational (domain) model that is suitable for formal analysis, both by simulation and by mathematical analysis. Moreover, within the *Integration and Projects* stream students learn how such domain models can be built in software systems in order to make them human-aware, so that they can provide support in a knowledgeable manner. In the *Integration and Projects* stream, students involve themselves with the analysis of questions and problems in areas of societal application and integrate the other ingredients of the curriculum into the design and implementation of solutions. These streams and their interactions are discussed in more detail in subsequent sections.

17.3 Computational Modeling Stream

Within the designed curriculum, the computational modeling stream plays a crucial role in integrating the ingredients from the exact and human sciences. As the curriculum is problem-oriented, specific societal problem areas are a point of departure, as discussed above. These are not just any areas, but are chosen within a specific scope or viewpoint, which indicates what they have in common. As discussed above, this demarcation relates to the possibility of providing support to humans by using devices in their environment with scientific knowledge from human sciences built in so that these devices have a justified understanding of the human processes considered. At the modeling level this view can be translated into a generic, unified type of overall system model, which Treur (2008) called a *reflective coupled human-environment system*.

Firstly, a brief sketch is given of this overall modeling perspective. Many applications of support systems in general can be viewed as coupled human-environment systems, where ‘coupled’ means mutually interacting. For the specific type of systems considered here, however, the coupling also occurs in a reflective form; see also Chap. 16.

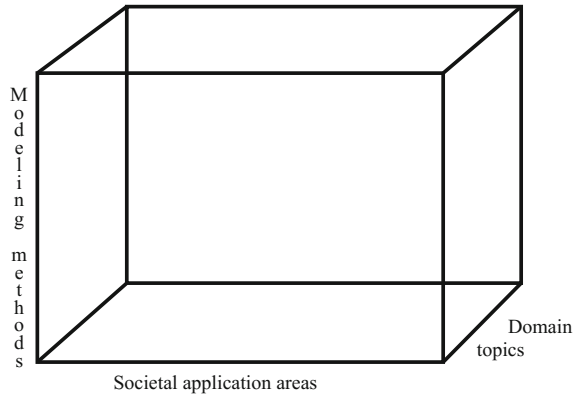
- On the one hand, coupling takes place as *interaction* between human and environment:
 - the environment receives information generated by the human, and
 - the human receives information generated by the environment.
- On the other hand, coupling at a *reflective level* takes place due to the fact that
 - in specific computational devices the environment has and maintains knowledge about the functioning of the human, the environment and their interaction, and
 - the human has and maintains knowledge about his or her own functioning, the environment, and their interaction

In order to realise applications according to the overall unified modeling perspective displayed in Fig. 17.2, a number of more specific ingredients for modeling both natural and artificial processes and their combinations are needed:

- (1) Computational *domain models* for human processes at the physiological, neurological, cognitive, affective and social levels;
- (2) *Integrative computational agent models* for software agents to support humans in their functioning, incorporating domain models with knowledge about human processes and methods for reasoning about them;
- (3) *Interaction models* for the interaction between software agent models and the environment, including sensor systems which can acquire information without having to bother the humans.

These ingredients can be obtained from areas such as Computational Science, Artificial Intelligence, Ambient Intelligence, and Informatics. In the curriculum,

Fig. 17.2 Three-dimensional space for applications



such ingredients are included in the Exact Sciences stream, and in the Computational Modeling stream students learn how to use and integrate these elements with domain knowledge as included in the Human Sciences stream (see Sect. 17.4). Examples of methods and techniques covered include: qualitative, logical, quantitative, numerical, and hybrid dynamical modeling; recursive modeling and model-based reasoning in software agent models using domain models; methods of analysis, assessment and intervention action generation. Given the specific motivations and backgrounds of the students at which this curriculum was to be aimed, much work was needed to develop course material in such a way that it fits well with these motivations and backgrounds. For example, much of the available literature on computational modeling is presented in a rather technical form, with examples often taken from engineering and physical sciences (Shiflet and Shiflet 2006). Such methods and techniques have been adapted to focus more on the human perspective. Moreover, the integration of domain models within (software) agent models is an area which is still under development, and developing course material for this has gone hand in hand with research. The same applies to the topic of model abstraction. More specifically, the following courses have been developed and are included in the computational modeling stream:

- *Introduction to Modeling and Simulation*

This course addresses the dynamical (domain) modeling of human processes, using elements of numerical, logical and hybrid computational modeling as described, for example, in Ashby (1952), Beer (1995a), Port and van Gelder (1995), Busemeyer and Diederich (2010), Shiflet and Shiflet (2006), Bosse et al. (2007) and Miller and Page (2007).

- *Integrative Modeling*

This course addresses the integration of domain models into models for supporting software applications. Here, elements from computational modeling, knowledge model and agent modeling are combined; see, for example, Bosse et al. 2011b, c.

- *Comparative modeling*

In this course, relationships between models are addressed, according to three dimensions of abstraction and inter-level relations as discussed in Bosse et al. (2010), Treur (2011a, b). The three abstraction dimensions addressed are the process abstraction dimension (see, for example, Sharpanskykh and Treur 2012a, b; Treur 2011a, b), the temporal abstraction dimension (e.g. Bosse et al. 2009) and the agent cluster abstraction dimension (e.g. Bosse et al. 2011, 2012; Sharpanskykh and Treur 2011).

- *Behaviour Dynamics in Social Networks*

This course addresses in more depth, and from a Network-Oriented Modeling perspective, the dynamics of cognitive, affective and social processes and the interaction between these processes. It also addresses analysis and validation of the network models. Much of the material of this book has been developed or used in this course.

- *Model-Based Intelligent Environments*

This course develops integrative modeling to more realistic overall intelligent environments.

17.4 The Human Sciences and Exact Sciences Streams

As mentioned in Sect. 17.3 above, in the curriculum ingredients are used from the Exact Sciences stream and in particular from areas such as Artificial Intelligence (e.g. modeling knowledge and reasoning), Ambient Intelligence (e.g. interaction with humans using sensor systems), and Informatics (e.g. human-computer interaction). More specifically, the Exact Sciences stream contributes courses such as:

- Logic and Sets, Intelligent Systems, Machine Learning, Evolutionary Computing (Artificial Intelligence)
- Pervasive Computing, Lab Human Ambience (Ambient Intelligence)
- Problem Solving, Introduction to Programming, Databases, Human-Computer Interaction, Multimedia authoring, Web Technology (Informatics)

For the Human Sciences stream, a large number of options is available from the existing curricula in human sciences such as Biomedical and Health Sciences, Psychology, and Social Sciences. Such courses may address more fundamental aspects of human processes, but also focus on limitations or shortcomings in functioning, for example due to specific disorders. Examples of courses for this stream are:

- Medical Physiology, Behaviour and Health (e.g. Widmaier et al. 2004)
- Introduction to Psychology and its Methods, Anxiety and mood disorders, Empirical methods (e.g. Gleitman et al. 2004; Ashcraft 2005; Nolen-Hoeksema 2005)
- Social Psychology, Text Analysis (e.g. Smith and Mackie 1999)

Within the curriculum, in a number of cases choices can be made by students for subjects or profiles they prefer. In this way they can create a specialisation, in themes such as mental health, sports, crime, and the elderly.

17.5 Integration and Projects

The Integration and Projects stream aims at integrating the different elements and streams in the curriculum, mostly in the form of project activities. A starting point for this stream is the integrative first year course Introduction Lifestyle Informatics, in which students design their first application. Later in the first year and in the second year there are integrative projects in which the knowledge obtained from the other streams thus far is integrated. In the third year, the Bachelor study ends with a larger integrative project. Similarly, in the fifth year the 2-year Master study ends with a larger integrative project which takes about half a year fulltime.

In the fourth year, the first year of the 2-year Master programme, the integrative course Human Ambience Innovation aims at getting an overview of the field, on the basis of the following three dimensions (see also Fig. 17.2):

- The modeling and implementation methods and techniques used;
- The domain knowledge from human sciences used;
- The societal application area in which a problem is addressed.

Each application can be mapped or projected on each of these axes, thus providing a triple-fold characterising of it. As a simple example, an application to support a depressed person through the Internet and their mobile phone, exploiting a causal model of how depression can develop, can be characterised as the following triple

<causal modeling and mobile Internet; psychological knowledge about depression; mental healthcare>

As another example, if a person's social environment is addressed by an application to avoid becoming socially isolated, for which a dynamical system model and Twitter are used, then the application can be characterised as:

<numerical dynamical system modeling and Twitter; knowledge about social interaction; mental healthcare>

17.6 Evaluation and Discussion

The contents of this chapter are mainly based on Treur (2013). In Treur (2007) a description is included of the designed curriculum in much more detail, including, for example, how it relates to the Dublin descriptors (e.g., Joint Quality Initiative Group 2004).

Higher education curriculum design is not a well-developed area of research; for an impression of different perspectives and meanings, see Fraser and Bosanquet (2006), Mäkinen and Annala (2010). Curricula can be developed for existing scientific disciplines; specific case studies have been reported in different domains, for example in automotive engineering and history (Shay 2011; Mears et al. 2011). However, curricula can also be designed for newly-developing disciplines or multidisciplinary areas, such as sustainability science (e.g. Michalcic et al. 2003). As both the scientific area and the curriculum are developing, these provide an extra level of challenge. The current chapter reports a curriculum design for such a multidisciplinary area. The curriculum as presented above was designed in a coherent manner according to a well-defined viewpoint, and was meant to contribute a serious innovation to the landscape of academic curricula. It does not look like many more common or traditional curricula.

A question raised by such a curriculum is whether it actually has an academic character. It seems clear that, when compared to any of its mono-disciplinary ingredients, it has less depth. However, the academic value is in the integration of the different scientific ingredients. In that sense, more depth is achieved, even in terms of the different mono-disciplinary ingredients. For example, software systems are designed in a manner that is justified by knowledge from human sciences, which has more depth than software systems that are only tested by users in relation to whether they appreciate the system. As another example, knowledge from human sciences is not only acquired in informal ways, but also in more formalized, in-depth forms based on computational models. In that sense, students in this curriculum achieve more depth than, say, psychology students who do not cover computational modeling.

Given the way in which it deviates substantially from known curricula, the process for implementing a curriculum as described above is certainly not straightforward. Although a first aim was to add this as a completely new curriculum to the spectrum of already-available curricula, after some time it turned out that the chances to get such a curriculum realised were considered politically much higher when it was to replace an existing curriculum. At this point the decision was made to implement this curriculum for the first 3 years as a replacement for the existing Bachelor course Artificial Intelligence (including adopting a new name for this Bachelor: Lifestyle Informatics), and for the fourth and fifth year as a specific profile (called Human Ambience) within the existing Master course Artificial Intelligence. Both are organised by the Department of Computer Science of the Faculty of Exact Sciences in cooperation with different faculties for human sciences.

In the initial phase of this whole process, in spring 2007 an investigation was conducted by a professional organisation to estimate the appreciation of such a course by candidate students. This study focused on penultimate-year high school students from several schools in the Netherlands, and had 1104 respondents (Hamstra 2007). This group of students could, after their next year choose to follow this curriculum. After a brief description of the curriculum (which was a Dutch variant of the text of the third paragraph included in Sect. 17.1 above), they gave an

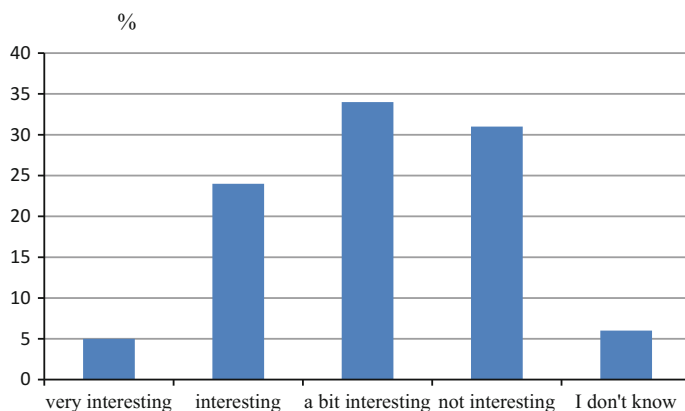


Fig. 17.3 Interests in the curriculum of 1104 students in the penultimate year of high school

answer to the question: *How far do you find this bachelor interesting, after reading the description?* Some of the results are depicted in Fig. 17.3. It turned out that a majority (63 %) of these students found such a study interesting, varying from a bit interesting (34 %) to interesting (24 %) and very interesting (5 %). Only 31 % of respondents found the study not interesting. One of the more specific positive outcomes was that, to a significant extent, the female students were the most positive ones. Only 25 % of these found the course not interesting, whereas 39 % of male students said the same.

In the meanwhile, the curriculum is functional for more than 5 years. In these years the number of students attracted by the new Lifestyle Informatics Bachelor course (up to 40 new students per year) was about double the number that were attracted by the Artificial Intelligence Bachelor it was replacing. Half of these students is female, whereas for the original AI bachelor this was less than 10 %. The percentages of students who successfully continue the course after the first and second year are high.

Regarding the question ‘Should I try this at home?’, the following can be said. On the one hand, be aware of some of the difficulties in implementing this, both in terms of content and politics. In this chapter they have been pointed out from time to time. On the other hand, be aware of the great opportunities such a curriculum offers in attracting new types of students; for example, it becomes possible to form groups of students that have an even balance of males and females, which is quite exceptional in an exact academic context. As highlighted in this chapter, integration through computational modeling is considered a crucial factor in order to obtain coherency in the curriculum. Drawing on the experiences described above, an important piece of advice would be to address that area very seriously. In contact with the author it may be possible to obtain dedicated course materials that have already been developed.

Acknowledgments The author is grateful to a large number of colleagues, who have, during discussions, deepened his insight into this area, and provided encouragement to go ahead with the challenging enterprise, among whom are Tibor Bosse, Frank van Harmelen, Mark Hoogendoorn, Johan Hoorn, Michel Klein, Peter-Paul van Maanen, Andre Spijkervet, Maarten van Steen, Gerrit van der Veer, Chris Verhoef, and Natalie van der Wal.

References

- R. Ashby, *Design for a Brain* (Chapman and Hall, London, 1952)
- M.H. Ashcraft, *Cognition* (Prentice Hall, Upper Saddle River, 2005)
- R.D. Beer, A dynamical systems perspective on agent-environment interactions. *Artif. Intell.* **72**, 173–215 (1995a)
- T. Bosse, C.M. Jonker, L. van der Meij, J. Treur, A language and environment for analysis of dynamics by simulation. *Int. J. Artif. Intell. Tools* **16**, 435–464 (2007)
- T. Bosse, C.M. Jonker, L. van der Meij, A. Sharpanskykh, J. Treur, Specification and verification of dynamics in agent models. *Int. J. Coop. Inf. Syst.* **18**, 167–193 (2009)
- T. Bosse, M. Hoogendoorn, M.C.A. Klein, J. Treur, A three-dimensional abstraction framework to compare multi-agent system models, in *Proceedings of the Second International Conference on Computational Collective Intelligence, ICCCI'10, Part I*. Lecture Notes in Artificial Intelligence, vol. 6421 (Springer, 2010), pp. 306–319
- T. Bosse, C.G. Gerritsen, M. Hoogendoorn, S.W. Jaffry, J. Treur, Agent-based versus population-based simulation of displacement of crime: a comparative study. *Web Intell. Agent Syst. J.* **9**, 147–160 (2011a)
- T. Bosse, M. Hoogendoorn, M.C.A. Klein, R.M. van Lambalgen, P.P. vanMaanen, J. Treur, Incorporating human aspects in ambient intelligence and smart environments, in *Handbook of Research on Ambient Intelligence and Smart Environments: Trends and Perspectives*, eds. by N.Y. Chong, F. Mastrogianni (IGI Global, 2011b), pp. 128–164
- T. Bosse, M. Hoogendoorn, M.C.A. Klein, J. Treur, An ambient agent model for monitoring and analysing dynamics of complex human behaviour. *J. Ambient Intell. Smart Environ.* **3**, 283–303 (2011c)
- T. Bosse, S.W. Jaffry, G. Siddiqui, J. Treur, Comparative analysis of agent-based and population-based modelling in epidemics and economics. *Multi-Agent Grid Syst. J.* **8**, 223–255 (2012)
- J.R. Busemeyer, J.R. Diederich, *Cognitive Modeling* (SAGE Publications, 2010)
- S.P. Fraser, A.M. Bosanquet, The curriculum? That's just a unit outline, isn't it? *Stud. High. Educ.* **31**, 269–284 (2006)
- H. Gleitman, A.J. Fridlund, D. Reisberg, *Psychology*, 6th edn. (Norton & Company Inc, New York, 2004)
- G. Hamstra et al., *Haalbaarheidsonderzoek Human Ambience* (Right Marktonderzoek en Advies B.V., 2007)
- Joint Quality Initiative Group, Shared 'Dublin' descriptors for short cycle, first cycle, second cycle and third cycle awards (2004). www.jointquality.org/content/descriptors/CompletesetDublinDescriptors.doc
- M. Mäkinen, J. Annala, Meanings behind curriculum development in higher education. *Prime* **4**, 1744–2494 (2010)
- L. Mears, M. Omar, T.R. Kurfess, Automotive engineering curriculum development: case study for Clemson University. *J. Intell. Manuf.* **22**, 693–708 (2011). doi:[10.1007/s10845-009-0329-z](https://doi.org/10.1007/s10845-009-0329-z)
- J.R. Michelic, J.C. Crittenden, M.J. Small, D.R. Shonnard, D.R. Hokanson, Q. Zhang, H. Chen, S.A. Sorby, V.U. James, J.W. Sutherland, J.L. Schnoor, Sustainability science and engineering: the emergence of a new metadiscipline. *Environ. Sci. Technol.* **37**, 5314–5324 (2003)

- J.H. Miller, S.E. Page, *Complex Adaptive Systems: An Introduction to Computational Models of Social Life* (Princeton Studies in Complexity, 2007)
- S. Nolen-Hoeksema, *Abnormal Psychology* (McGraw-Hill, 2005)
- R. Port, T. van Gelder, *Mind as Motion: Explorations in the Dynamics of Cognition* (MIT/Bradford, 1995)
- A. Sharpanskykh, J. Treur, Group Abstraction for Large-Scale Agent-Based Social Diffusion Models, in *Proceedings of the Third International Conference on Social Computing, SocialCom'11*, eds. by J. Zhan, M. Pantic, A. Vinciarelli (IEEE Computer Society Press, 2011), pp. 830–837
- A. Sharpanskykh, J. Treur, Abstraction relations between internal and behavioural agent models for collective decision making. *Web Intell. Agent Syst. J.* **10**, 465–484 (2012a)
- A. Sharpanskykh, J. Treur, An ambient agent architecture exploiting automated cognitive analysis. *J. Ambient Intell. Humaniz. Comput.* **3**, 219–237 (2012b)
- S. Shay, Curriculum formation: a case study from history. *Stud. High. Educ.* **36**, 315–329 (2011)
- A.B. Shiflet, G.W. Shiflet, *Introduction to Computational Science: Modeling and Simulation for the Sciences* (Princeton University Press, 2006)
- E.R. Smith, T.M. Mackie, *Social Psychology* (Worth Publishers, New York, 1999)
- J. Treur, *Bachelor Study Human Ambience (in Dutch). Report* (VU University Amsterdam, Department of Computer Science, Amsterdam, 2007), p. 83
- J. Treur, On human aspects in ambient intelligence, in *Proceedings of the First International Workshop on Human Aspects in Ambient Intelligence, HAI'07*. Published in: *Communications in Computer and Information Science (CCIS)*, vol. 11 (Springer, 2008), pp. 262–267
- J. Treur, On the use of reduction relations to relate different types of agent models. *Web Intell. Agent Syst. J.* **9**, 81–95 (2011a)
- J. Treur, Specification of interlevel relations for agent models in multiple abstraction dimensions, in *Proceedings of the 24th International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems, IEA/AIE'11, Part II*, ed. by K.G. Mehrotra et al. *Lecture Notes in Artificial Intelligence*, vol. 6704 (Springer, 2011b), pp. 542–555 (Extended version in: *International Journal of Modeling, Simulation, and Scientific Computing*, vol. 4(1), 1250026, 2013, pp. 1–27)
- J. Treur, Designing a problem-oriented multi-disciplinary academic curriculum: integrating biomedical, psychological, and social sciences with ambient intelligence, artificial intelligence and informatics, in *Proceedings of the Third World Conference on Learning, Teaching and Educational Leadership, WCLTA'12*, ed. by H.F. Odabasi. *Procedia Social and Behavioral Sciences*, vol. 93 (Elsevier, 2013), pp. 258–265
- E.P. Widmaier, H. Raff, K.T. Strang, *Vander, Sherman en Luciano's Human Physiology* (MacGraw Hill, 2004)

Part VI

Network-Oriented Modeling: Discussion

Chapter 18

On the Use of Network-Oriented Modeling

A Discussion

Abstract This chapter is a discussion in which some of the main issues addressed in the book are briefly reviewed. In particular, Network-Oriented Modeling based on adaptive temporal-causal networks is discussed and how generic and applicable it is as a modeling approach and as a computational paradigm.

18.1 Introduction

This book started in Chap. 1 by a review of traditionally used means to address the complexity of individual and social human processes. These means often concern assumptions on separation and isolation of parts of processes. Due to the shortcomings of these assumptions, over time they have often led to strong debates. Many human processes involve sub-processes running simultaneously in parallel, thereby intensely interacting in cyclic manners. This offers an important challenge to be addressed, and it was recognized that a modeling perspective is needed that addresses such intense cyclic interactions and their dynamics. A Network-Oriented Modeling perspective was proposed here as an alternative way to address complexity. Using this perspective, different elements of a process can be distinguished, but it does not separate or isolate them. Instead it emphasizes and explicitly models how they run and interact simultaneously. By incorporating a temporal dimension, it is modeled how they can have intense and circular causal interaction, and the timing of such processes can be modeled.

18.2 Network-Oriented Modeling

Although the notion of network itself and its use in different contexts can be traced back to the years 1930–1960 (see Chap. 1, Sect. 1.4), the notion of Network-Oriented Modeling as a modeling approach (also indicated by NOM) can be found only in more recent literature, and only for specific domains. More specifically, this notion is used

in different forms in the context of modeling organisations and social systems (e.g., Elzas 1985; Chung et al. 2003; Naudé et al. 2008), of modeling metabolic processes (e.g., Cottret and Jourdan 2010), and of modeling electromagnetic systems (e.g., Russer and Cangellaris 2001; Felsen et al. 2002, 2009). The Network-Oriented Modeling approaches put forward in this literature are specific for the domains addressed, respectively social systems, metabolic processes and electromagnetic systems. An interesting challenge is to achieve unification of such Network-Oriented Modeling methods. The Network-Oriented Modeling approach described in this book was developed with the domain of mental and social human processes in mind (but also with inspiration from modeling metabolic processes within bacteria; e.g., Jonker et al. (2002, 2008)), thus unifying at least both individual human processes and social processes, as has been illustrated by many example models in this book. However, the scope of applicability is much wider, as discussed in some more detail in Sect. 18.3.

The Network-Oriented Modeling approach presented in this book uses adaptive temporal-causal networks as a vehicle. The temporal perspective allows to model the dynamics of the interaction processes within networks and of networks well. A conceptual representation of a model represents in a declarative manner *states* and *connections* between them. States have (activation) levels that vary over time. The connections stand for (causal) impacts of states on each other. Furthermore, the notion of *weight of a connection* is used to be able to express differences in strengths of impact. Moreover, *combination functions* are used to express how to *aggregate multiple causal impacts* on a state. Within adaptive networks also these weights can vary over time. Finally, the notion of *speed factor* expresses the speed of change of a state and is used to model timing of processes.

18.3 Genericity of a Network-Oriented Modeling Approach

In this section it is discussed how generic the presented Network-Oriented Modeling approach is. More specifically, it is discussed how temporal-causal networks subsume smooth continuous dynamical systems, discrete dynamical systems and computational processes more in general.

Network-Oriented Modeling and Continuous Dynamical Systems

In Chap. 2, Sect. 2.9 it has been discussed that any smooth continuous dynamical system (which by definition is a state-determined system) can be modeled as a temporal-causal network model, by choosing suitable parameters such as connection weights, speed factors and combination functions. In this sense this Network-Oriented Modeling approach is as general as dynamic modeling approaches put forward, for example, in Ashby (1960), Forrester (1973, 1987), Thelen and Smith (1994), Port and van Gelder (1995), van Gelder and Port (1995), Beer (2000), Kelso (1995), van Gelder (1998), and neural network approaches such as described,

for example in Grossberg (1969), Hopfield (1982, 1984), Hirsch (1989) and Funahashi and Nakamura (1993). This indicates that using this Network-Oriented Modeling approach does not limit the scope of the modeling.

Network-Oriented Modeling and Discrete Dynamical Systems

The numerical representations of temporal-causal network models can be used to model continuous dynamical systems. But they can also be used to model discrete binary processes based on values 0 or 1 for the states. To this end, set time step $\Delta t = 1$, speed factor $\eta_Y = 1$ for all states Y , connection weight $\omega_{X,Y} = 1$ for all states X and Y with a connection from X to Y , and assume that all combination functions $c_Y(\dots)$ only generate values 0 or 1, when applied to values 0 or 1. Then the difference equation for a state Y becomes

$$Y(t+1) = Y(t) + [c_Y(X_1(t), \dots, X_k(t)) - Y(t)]$$

which simply can be rewritten as:

$$Y(t+1) = c_Y(X_1(t), \dots, X_k(t))$$

This takes the form of a general evolution or transition rule for a discrete dynamical system of which the (overall) states are defined as vectors $(X_1(t), \dots, X_k(t))$ with values 0 or 1, and transitions of overall states are defined as

$$(X_1(t+1), \dots, X_k(t+1)) = (c_{X_1}(X_1(t), \dots, X_k(t)), \dots, c_{X_k}(X_1(t), \dots, X_k(t)))$$

or in vector notation $\mathbf{X}$ with $\mathbf{X}(t) = (X_1(t), \dots, X_k(t))$:

$$\mathbf{X}(t+1) = \mathbf{c}(\mathbf{X}(t))$$

where for $\mathbf{V} = (V_1, \dots, V_k)$ it is defined $\mathbf{c}(\mathbf{V}) = (c_{X_1}(\mathbf{V}), \dots, c_{X_k}(\mathbf{V}))$.

This shows how the Network-Oriented Modeling approach based on temporal-causal networks subsumes modeling by discrete dynamical systems. Note that the above approach abstracts from the temporal aspect by setting Δt and all speed factors 1. However, also timed variants of discrete dynamical systems can be covered.

Network-Oriented Modeling and Computational Processes

Any real implemented computational process in principle is a deterministic smooth continuous process of a state-determined system in the physical world. Therefore it could be claimed that the temporal-causal network modeling approach in theory covers all computational processes. Within theoretical analyses often variants of transition systems or finite state machines are used as universal ways to specify computational processes. Conceptually such types of representations of (state) transitions can easily be related to causal relations as considered in the temporal-causal network modeling approach. In more detail, the format for discrete dynamical systems described above as a special case can be used to model transition systems or finite state machines within the temporal-causal network modeling

approach: by defining $c(X) = Y$ if and only if within a finite state machine or transition system there is a transition from the overall state represented as X to the overall state represented as Y . This also provides support for the theoretical claim that computational processes can be covered by the temporal-causal network modeling approach. However, to support such a general claim for any specific practical computational paradigm could be a nontrivial challenge. For example, although perhaps theoretically possible, to obtain a temporal-causal network representation for a computational process described in some procedural (parallel) programming language, in practice may require some effort. This may be similar to transformations of procedural specifications into other types of declarative representations, for example, into (temporal) logical or functional formats.

18.4 Applicability of Network-Oriented Modeling

The Network-Oriented Modeling approach has turned out useful in particular for computational modeling in a multidisciplinary context. Moreover, network models as obtained can form a solid basis to develop smart applications.

Applicability for Modeling in a Multidisciplinary Context

As discussed in Chaps. 1 and 2 the temporal-causal network modeling approach used here makes it easy to take into account theories and findings about dynamics of processes from any scientific discipline, as commonly such processes are described in terms of causal relations.

In particular, this applies to complex brain processes known from Cognitive, Affective and Social Neuroscience, which, for example, often involve dynamics based on interrelating cycles. Also recall the quotation of Phelps in Chap. 1, Sect. 1.2: ‘Adding the complexity of emotion to the study of cognition can be daunting, but investigations of the neural mechanisms underlying these behaviours can help clarify the structure and mechanisms’ (Phelps 2006, pp. 46–47). A Network-Oriented Modeling approach enables to address in an integrative manner complex cognitive, affective and social phenomena such as dynamics by or of social interaction, the integration of emotions within cognitive processes, internal simulation of external processes, mirroring of mental processes of others, and Hebbian learning; e.g., Hebb (1949), Gerstner and Kistler (2002), Keyzers and Perrett (2004) and Keyzers and Gazzola (2014). It also has been discussed in Chap. 1 how a Network-Oriented Modeling approach relates to perspectives in Philosophy of Mind (e.g., Kim 1996), in particular to (causal) networks of mental states. Furthermore, it has been discussed in Chap. 1 how the approach relates to the philosophical perspective on dynamics in the physical world that is indicated as the clockwork universe; e.g., Descartes (1634) and Laplace (1825). In an abstract sense this perspective relates to the notion of state-determined system; e.g., Ashby (1960).

For processes in a social context, social phenomena such as shared understanding and collective power show how bridges between individual persons are constructed. The behaviour of each person is based on internal states such as goals, emotions and beliefs. Therefore from a naïve viewpoint such sharedness and collectiveness would be considered as very improbable. But specific mechanisms do their work in tuning the individual mental processes to each other, mostly in an unconscious manner, and lead to the emergence of shared mental states and collective behaviour. Knowledge about these mechanisms from Social Neuroscience can be exploited to model corresponding computational mechanisms. It has been discussed in Chap. 7 how from a neuroscientific perspective, mirror neurons and internal simulation are core mechanisms for this.

From the applications to model complex phenomena by a Network-Oriented Modeling approach, within the book models for the following complex phenomena in a multidisciplinary context have been discussed:

- Embodiment, as-if body loops, mindfulness (Chap. 3)
- Imagination, visualisation and dreaming as internal simulation (Chap. 4)
- Mirroring of other minds (Chap. 7)
- Integration of affective and cognitive processes (Chaps. 3, 6, 7, 10)
- Fear extinction learning (Chap. 5)
- Emotions as a basis for rationality (Chap. 6)
- Empathic understanding (Chaps. 7, 9)
- Emergence of shared understanding and collective action (Chap. 7)
- Group processes and crowd behavior (Chap. 7)
- Prior and retrospective ownership of actions (Chap. 8)
- Social contagion (Chaps. 7, 11)
- Social responsiveness (Chap. 9)
- Joint decisions (Chap. 10)
- Social network evolution (Chap. 11).

Applicability for the development of Smart Applications

The topics addressed have a number of possible applications. An example of such an application is to analyse the spread of a healthy or unhealthy lifestyle in society. Another example is to analyse crowd behaviour in emergency situations. A wider area of application, as discussed in Chap. 16, addresses smart applications in the context of Ambient Intelligence or socio-technical systems that consist of humans and devices, such as smartphones, and use of social media. For such applications, in addition to analysis of the relevant processes, also for the support side the design of these devices and media can be an important aim. This may concern, for example, safe evacuation in an emergency situation or strengthening development of a healthy lifestyle. Other application areas may address, for example, support and mediation in collective decision making and avoiding or resolving conflicts that may develop. The Network-Oriented Modeling approach as presented makes modeling complex human and social processes more manageable, and extends the range of what is possible. To facilitate applications, dedicated software is available

supporting the design of network models in a conceptual manner, automatically transforming them into an executable format and performing simulation experiments.

18.5 Finally

Summarizing, the Network-Oriented Modeling approach based on temporal-causal networks as described here, provides a complex systems modeling approach that enables a modeler to design high level conceptual model representations in the form of cyclic graphs (or connection matrices), which can be systematically transformed in an automated manner into executable numerical representations that can be used to perform simulation experiments. The modeling approach makes it easy to take into account on the one hand theories and findings from any domain from, for example, biological, psychological, neurological or social sciences, as such theories and findings are often formulated in terms of causal relations. This applies, among others, to mental processes based on complex brain networks, which, for example, often involve dynamics based on interrelating and adaptive cycles, but equally well it applies to social networks and their adaptive dynamics. This enables to address complex adaptive phenomena such as the integration of emotions within all kinds of cognitive processes, of internal simulation and mirroring of mental processes of others, and dynamic social interaction patterns. By using temporal-causal relations from those domains as a main vehicle and structure for network models, the obtained network models get a strong relation to the large body of empirically founded knowledge from the Neurosciences and Social Sciences. This makes them scientifically justifiable to an extent that is not attainable for black box models which lack such a relation.

References

- W.R. Ashby, *Design for a Brain*, 1st edn. 1952 (Chapman and Hall, London (second extended edition), 1960)
- R.D. Beer, Dynamical approaches to cognitive science. *Trends Cogn. Sci.* **4**, 91–99 (2000)
- B. Chung, H. Choi, S. Kim, Workflow-enabled internet service delivery for a variety of access networks. The 7th Asia-Pacific network operations and management symposium (APNOMS, 2003)
- L. Cottret, F. Jourdan, Graph methods for the investigation of metabolic networks in parasitology. *Parasitology* **137**(09), 1393–1407 (2010)
- R. Descartes, *The World or Treatise on Light* (translated version by M.S. Mahoney) (1634). <http://www.princeton.edu/~hos/mike/texts/descartes/world/world.htm>
- M.S. Elzas, Organizational structures for facilitating process innovation, in *Real Time Control of Large Scale Systems* (Springer Berlin Heidelberg, 1985), pp. 151–163

- L.B. Felsen, M. Mongiardo, P. Russer, Electromagnetic field representations and computations in complex structures I: complexity architecture and generalized network formulation. *Int. J. Numer. Model. Electron. Netw. Dev. Fields* **15**(1), 93–107 (2002)
- L.B. Felsen, M. Mongiardo, P. Russer, *Electromagnetic Field Computation by Network Methods* (Springer Science & Business Media, 2009)
- J.W. Forrester, *World Dynamics*, 2nd edn. (Pegasus Communications, Waltham, MA, 1973). pp. 144
- J.W. Forrester, Lessons from system dynamics modeling. *Syst. Dyn. Rev.* **3**(2), 136–149 (1987)
- K. Funahashi, Y. Nakamura, Approximation of dynamical systems by continuous time recurrent neural networks. *Neural Netw.* **6**, 801–806 (1993)
- W. Gerstner, W.M. Kistler, Mathematical formulations of Hebbian learning. *Biol. Cybern.* **87**, 404–415 (2002)
- S. Grossberg, On learning and energy-entropy dependence in recurrent and nonrecurrent signed networks. *J. Stat. Phys.* **1**, 319–350 (1969)
- D. Hebb, *The Organisation of Behavior* (Wiley, 1949)
- M. Hirsch, Convergent activation dynamics in continuous-time networks. *Neural Netw.* **2**, 331–349 (1989)
- J.J. Hopfield, Neural networks and physical systems with emergent collective computational properties. *Proc. Nat. Acad. Sci. (USA)* **79**, 2554–2558 (1982)
- J.J. Hopfield, Neurons with graded response have collective computational properties like those of two-state neurons. *Proc. Nat. Acad. Sci. (USA)* **81**, 3088–3092 (1984)
- C.M. Jonker, J.L. Snoep, J. Treur, H.V. Westerhoff, W.C.A. Wijngaards, Putting intentions into cell biochemistry: an artificial intelligence perspective. *J. Theor. Biol.* **214**, 105–134 (2002)
- C.M. Jonker, J.L. Snoep, J. Treur, H.V. Westerhoff, W.C.A. Wijngaards, BDI-modelling of complex intracellular dynamics. *J. Theor. Biol.* **251**(2008), 1–23 (2008)
- J.A.S. Kelso, *Dynamic patterns: the self-organization of brain and behavior* (MIT Press, Cambridge, 1995)
- C. Keysers, V. Gazzola, Hebbian learning and predictive mirror neurons for actions, sensations and emotions. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **369**, 20130175 (2014)
- C. Keysers, D.I. Perrett, Demystifying social cognition: a Hebbian perspective. *Trends in Cog. Sci.* **8**, 501–507 (2004)
- J. Kim, *Philosophy of Mind* (Westview Press 1996)
- P.S. Laplace, *Philosophical Essays on Probabilities* (translated by A.I. Dale from the 5th French edition of 1825) (Springer, New York, 1825)
- A. Naudé, D. Le Maitre, T. de Jong, G.F.G. Mans, W. Hugo, Modelling of spatially complex human-ecosystem, rural-urban and rich-poor interactions (2008). https://www.researchgate.net/profile/Tom_De_jong/publication/30511313_Modelling_of_spatially_complex_human-ecosystem_rural-urban_and_rich-poor_interactions/links/02e7e534d3e9a47836000000.pdf
- E.A. Phelps, Emotion and cognition: insights from studies of the human amygdala. *Annu. Rev. Psychol.* **57**, 27–53 (2006)
- R.F. Port, T. van Gelder, *Mind as motion: explorations in the dynamics of cognition* (MIT Press, Cambridge, MA, 1995)
- P. Russer, A.C. Cangellaris, Network oriented modeling, complexity reduction and system identification techniques for electromagnetic systems, in *Proceedings. 4th International. Workshop on Computational Electromagnetics in the Time-Domain: TLM/FDTD and Related Techniques* (2001), pp. 105–122
- E. Thelen, L. Smith, *A dynamic systems approach to the development of cognition and action* (MIT Press, Cambridge, 1994)
- T. van Gelder, The dynamical hypothesis in cognitive science. *Behav. Brain Sci.* **21**(1998), 615–665 (1998)
- T. van Gelder and R.F. Port (1995). It's about time: an overview of the dynamical approach to cognition, in *Mind as Motion: Explorations in the Dynamics of Cognition* (MIT Press, Cambridge, MA, 1995), pp. 1–43

Index

A

Achievement property, 368, 372, 373, 387
Action, 5, 25, 28, 42, 116, 127, 184, 187, 198, 204, 209, 220, 226, 238, 279, 429, 453, 469
Action ownership, 25, 197, 198, 205, 209, 210, 212, 213, 214, 217, 218, 229, 231, 276, 278, 279, 282, 491
Activation level, 23, 128, 190, 353
Actualisation, 423, 431, 443
Adaptation model, 470
Adapting parameters as states, 93
Adaptive network, 21, 97, 263, 298, 304, 317, 324
Adaptive temporal-causal network, 24–26, 120, 142, 159, 265
Advanced logistic sum function, 71, 131
Advanced product-based combination function, 79
Affective-focused empathy, 196
Affective-focused ToM, 196
Affective neuroscience, 35
Affective preparation loop, 189
Affective type of understanding, 193
Aggregated impact, 21, 59–61, 67, 71, 77, 79, 80, 303, 346, 396, 459
Ambient intelligence, 28, 463, 465, 470, 476, 491
Amygdala, 8, 108, 136, 152, 159
Analysis model, 470
Antecedent-focused regulation, 111, 145, 240
Anterior cingulate cortex, 109, 152, 186, 228
Anticipatory state property, 433, 453
Applicability, scope of, 22, 39, 85, 488
As-if body loop, 6, 116, 158, 160, 188, 241, 271
Associated emotions, 23, 26, 138, 154, 200, 282
Autism spectrum disorder, ASD, 25, 236

Automated checking of dynamic properties, 387

Avoiding, 111, 257, 474

B

Behaviour, 23, 25, 41, 114, 175, 184, 190, 204, 210, 237, 249, 265, 299, 343, 349, 394, 435, 465, 469, 479, 490
Belief, 8, 26, 114, 115, 183, 194, 230, 374, 466
Body, 4, 5, 109, 185, 193, 248, 261
Body loop, 11, 108, 128, 159, 187, 244, 401
Body state, 5, 144, 188, 239, 386
Brain, 5–7, 13, 18, 37, 41, 42, 99, 108, 178
Bridge connection, 318

C

Causal factorisation principle, inward, 57
Causal factorisation principle, outward, 57
Causal graph, 18, 36, 423
Causality, 11, 240
Causal modeling, 36
Causal relation, 9, 18, 35, 39, 85, 245, 349, 440, 460, 492
Causal role, 12, 429, 435
Causal simulation, 36
Central monitoring theory, 211
Characteristics, personal, 25, 197
Clockwork universe, 99, 490
Cluster, 292, 293, 317, 479
Cognition, 4, 7, 18, 107
Cognition vs emotion, 7
Cognitive, 4, 7, 23, 28, 35, 42, 88, 105, 116, 183, 193, 200, 205, 216, 240, 270, 390, 423, 470, 490
Cognitive-affective interaction, 42, 121
Cognitive-affective loop, 190
Cognitive-focused empathy, 196
Cognitive-focused ToM, 196

Cognitive neuroscience, 37
 Cognitive science, 27, 421, 423, 464
 Cognitive type of understanding, 193
 Collective, 24, 178, 184, 197, 204, 491
 Collective action, 183, 185, 197, 204
 Combination function, 19, 22, 38, 46, 49, 59, 64, 71, 75, 81, 88, 94, 129, 314, 329
 Comparator model, 211
 Complementary product-based combination function, 70
 Complexity, 3, 4, 11, 14, 27, 44, 490
 Complex system, 21, 28, 492
 Computational methods, 37, 39, 99
 Computational model, 28, 114, 201, 230, 264, 356, 463, 475, 481
 Conceptual representation, 19, 38, 39, 47, 58, 59, 64, 99, 106, 115, 162, 300, 350, 363, 382
 Conjunction, 243
 Connection, 4, 13, 18, 21, 24, 40, 53, 60, 75, 83, 106, 111, 115, 137, 145, 150, 159
 Connection weight, 22, 49, 92, 112, 274, 300, 313, 332, 357, 395, 488
 Contagion principle, 9
 Control state, 23, 133, 145, 151, 249
 Coupled reflective system, 464, 468
 Cyclic causal network, 13, 36

D

Decision making, 7, 8, 22, 114, 157, 178, 199, 269, 281, 491
 Declarative modeling, 38, 99
 Decreasing, 26, 306, 313, 324, 375
 Desire, 8, 24, 28, 183, 385, 386, 423, 435, 448
 Deterministic world, 159, 166, 169
 Difference equation, first-order, 59, 63, 81, 90, 274
 Differential equation, first-order, 58, 86, 88, 92, 129, 336, 444, 446, 460
 Directed acyclic graph, DAG, 36
 Disjunction, 440, 441
 Domain model, 469, 470, 476, 478
 Dream episode, 23, 128, 138, 145, 154
 Dreaming, 10, 24, 125, 132, 141
 Dynamical system, 19, 85, 92, 212, 446, 458, 488, 489
 Dynamic property, 146, 215, 244, 350, 353, 358, 368, 372, 386
 Dynamic property expression, 366, 368
 Dynamics, 4, 12, 14, 16–18, 23, 27, 36, 37, 64, 99, 202, 203, 296, 315, 325, 390, 421, 432, 433, 460, 490

E

Economic choice, 159
 Educational perspective, 23
 Effectiveness rate, 164
 Emergence, emergent, 9, 24, 184, 197, 204, 350
 Emotion, 4, 6–8, 42, 77, 105, 106, 111, 112, 114–116, 125, 127, 157, 185, 187, 192, 198, 205, 236, 240, 245, 249, 263, 279, 469
 Emotion contagion, 119, 120
 Emotion integration, 25, 236, 244
 Emotion regulation, 10, 23, 25, 42, 106, 107, 111–114, 120, 121, 126–128, 137, 138, 142, 145, 147, 150, 153, 154, 235, 236, 239, 240, 244, 263, 264, 343
 Emotion-related valuing, 26, 178, 281
 Empathic understanding, 25, 184, 196, 205, 270, 281
 Empathy, 25, 195, 241
 Enhanced sensory processing sensitivity, 114, 240, 245, 249, 262
 Equilibrium, 6, 109, 203, 324, 328, 330, 332, 336, 343, 379, 382, 383
 Equilibrium state property, 379
 Error function, 398, 400, 404, 417
 Error measure, 398
 Evolutionary function of dreaming, 127
 Evolving network, 312
 Evolving social interactions, 26
 Exchange of potentialities by interaction, 435
 Execution state, 50, 56, 84, 168, 173, 300
 Exhaustive search, 383, 406, 412
 Existential quantifier, 369, 370, 387
 Expression, 5, 107, 186, 242, 249, 324, 328, 335, 350, 359, 362, 387
 Extinction rate, 94, 97, 260

F

Fear extinction learning, 10, 121, 125, 145, 153
 Feeling, 6, 7, 107, 109, 115, 128, 133, 145, 150, 153, 157, 192, 199, 240, 271, 279, 400
 Force, 3, 424, 435, 453, 456, 460
 Force of motion, 434
 Fulfilment, 397, 431
 Full empathy, 196
 Functional role, 99

G

Ghost, 422, 428, 436, 438, 442, 455
 Global workspace theory of consciousness, 154
 Gradient descent, 406, 407, 410, 412
 Graph, 40, 49, 55, 152, 252, 261, 332

Graph representation, conceptual, 88
 Grounding property, 376

H

Hebbian learning, 145, 148, 158, 167, 237, 260, 265, 324, 336, 347
 Higher order potentiality, 424
 Hill climbing, 410
 Hippocampus, 137, 152, 198
 Homophily principle, 42, 298, 305, 306, 317, 324
 Human-aware, 28, 476

I

Identity function, 49, 69, 314
 Impact, 12, 20, 48, 69, 80, 303
 Implication, 18, 441
 Increasing, 71, 323, 337, 459
 Individual, 8, 18, 195, 199, 204, 257, 474, 491
 Individual processes vs collective processes, 4
 Inhibiting link, 128, 143
 Insula, 109
 Insular cortex, 109
 Integrative modeling, 470, 479
 Interaction, 4, 7, 9, 11, 13–15, 18, 25, 40, 42, 66, 99, 114, 153, 236, 243, 263, 286, 298, 313, 435, 470, 476, 488
 Interaction connects principle, 298, 317
 Internal simulation, 8, 10, 24, 26, 125, 137, 144, 153, 160, 198, 210, 269, 492

J

Joint decision, 25, 197, 269, 270, 279, 281
 Joint decision making, 25, 230, 270, 281

L

Laws of nature, 17
 Learning rate, 168
 Learning social responses, 25
 Limit cycle, 13, 328, 345, 347, 379
 Limit cycle property, 385
 Linear momentum, 434
 Living force, 434

M

Maintenance property, 379
 Mathematical analysis, 59, 298, 324, 325, 347, 476
 Matrix representation, conceptual, 47, 49, 56, 129
 Max-based combination function, 70
 Medial prefrontal cortex, 136
 Mental causation, 432, 434, 435, 442, 443
 Metabolic process, 15, 488

Min-based combination function, 69, 77
 Mind, 5, 12, 18, 423
 Mindreading, 6
 Mind vs body, 4
 Mirror neuron, discovery, 8, 25, 119, 186, 199, 237, 265
 Mirror neurons, 8, 24, 26, 42, 119, 183–189, 192, 199, 203, 228, 229, 235–238, 241, 243, 259, 263, 265, 269–270, 273, 277, 281, 491
 Monotonic, 76, 80
 Monotonicity property, 376
 More becomes more principle, 26, 298, 313, 317
 Motive force, 434, 454
 Multidisciplinary curriculum, 475
 Multiple causal impacts, 20, 37
 Multiple impacts, 20
 Multi-realisability, 442, 455
 Mutually acknowledged, 25, 281
 Mutually acknowledged empathic understanding, 25

N

Naming option, 51
 Negation, 360
 Network, 14, 15, 18, 20, 26, 28, 40, 47, 58, 68, 90, 142, 158, 184, 210, 230, 243, 281, 285, 294, 296, 298, 303, 313, 324, 470, 489, 492
 Network analysis, 39, 82
 Network-Oriented Modeling, 36, 39, 45, 184, 323, 487, 491, 492
 Neuroeconomics, 159
 Neuroscience, 4, 8, 39, 41, 120, 193, 281, 490, 492
 Node degree, 289
 Non-adaptive processes vs adaptive processes, 4
 Nonnegative, 79, 80, 116, 294
 Numerical-logical representation, 350, 351, 358, 363, 369, 370, 389
 Numerical-logical representation of a dynamic property expression, 368
 Numerical representation, 19, 39, 58, 59, 64, 81, 110, 129, 142, 216, 286, 349, 364, 489, 492
 Numerical state relation representation, 369
 Numerical term representation, 365, 366

O

Ordering property, 376
 Ownership, 25, 197, 198, 205, 209, 210, 212–214, 217, 219, 220, 222, 223, 225–227, 229, 230, 273, 277, 279, 282, 491

P

- Pain example, 12
- Parameter, 19–22, 26, 27, 38, 49, 59, 69, 71, 75, 76, 89, 92, 93, 118, 129, 131, 138, 145, 168, 202, 203, 221, 223, 224, 226, 227, 243, 252, 257, 260, 264, 274, 275, 292, 302, 323, 324, 328, 329, 335, 350, 352, 353, 357, 372, 383, 384, 389, 393, 395–397, 400, 402, 404, 406, 407, 410, 412, 413, 417, 488
- Parameter estimation, 27, 394, 398, 400, 403, 407, 412, 417
- Parameter search, 372, 397, 403, 405
- Parameter tuning, 19, 383, 403
- Peak property, 377, 378
- Personalisation of a model, 351, 389
- Person-external understanding, 195
- Person-internal understanding, 196
- Pervasive computing, 28, 463, 479
- Philosophical perspective, 99, 490
- Physical realism, 28, 421, 422
- Plasticity, 42, 237
- Potentiality, 422–424, 427, 429, 430, 432–435, 438, 439, 441–445, 447, 450, 451, 453, 455, 457, 460
- Power, 17, 24, 42, 193, 197, 204, 290, 434, 491
- Power law, 289–291, 312, 313
- Prediction, 6, 10, 11, 25, 85, 118, 127, 138, 144, 160, 163, 164, 178, 197, 198, 209, 210, 212, 214, 220, 222–227, 230, 271, 274, 317, 401, 404, 428, 432, 438
- Prefrontal cortex, 112, 136, 152, 159, 198, 229
- Preparation state, 10, 50, 56, 70, 83, 84, 110, 118, 119, 127, 128, 131, 137, 138, 144–146, 150, 153, 160, 162, 165, 166, 168, 171, 186, 187, 189, 190, 192, 198, 199, 201, 205, 214, 216, 220, 221, 223–226, 228, 237, 239, 248, 271, 273, 276, 343, 352, 353, 374, 375, 377, 380
- Preservation of zero impact, 77, 80
- Prior ownership, 197, 209, 214, 217, 228, 230
- Product-based combination function, 70, 76

Q

- Quantity of motion, 423, 433, 453

R

- Random gradient descent, 394, 410, 411, 416
- Random network, 40, 286, 288, 289, 317
- Rational, 7, 8, 24, 118, 157–159, 169, 175, 176, 257

- Rationality measure, 159, 175, 178
- Realisation relation, 442, 444, 455, 460
- Realiser, 422, 442
- Realisers for derivatives, 449
- Relational specification, 374, 423, 439, 460
- Representation property, 385, 386
- Representation relation, backward, 374, 386
- Representation relation, forward, 374, 385
- Representation transformation, model, 58, 59
- Requirement, 16, 350, 351, 353, 354, 373, 394–398, 400, 404, 417
- Response-focused regulation, 111, 126, 145
- Retrospective ownership, 25, 197, 210, 213–215, 218, 227, 229, 230, 282

S

- Scaled sum function, 67, 77, 302, 311, 325, 333, 335, 336
- Scaled sum with interaction terms, 76
- Scale-free network, 40, 286, 289, 292, 317
- Scenario, 24, 25, 27, 92, 126, 128, 131, 133–136, 138, 142, 143, 146, 148, 150, 151, 154, 163, 166, 168, 172, 175, 176, 210, 214, 219–221, 223–227, 230, 257, 260, 262, 263, 270, 278, 281, 350, 352, 398, 440
- Schizophrenia, 25, 209, 210, 224, 230, 231
- Science of dynamics, 434
- Self-other distinction, 25, 187, 195, 236, 238, 241, 243, 244, 248, 249, 254, 263, 271, 277
- Sensor state, 50, 51, 53, 56, 57, 109, 112, 116, 119, 120, 164, 165, 188, 189, 192, 213, 216, 218, 243, 244, 248, 250, 251, 276, 278, 300–302, 330, 333, 343, 375, 377, 401, 402
- Sensory representation, 5, 10, 23, 24, 50, 51, 56, 83, 107, 108, 112, 116, 119, 125, 127–129, 131–135, 137, 138, 141, 143–148, 150–152, 154, 158, 160, 164–168, 171, 187, 189, 192, 193, 198, 200, 204, 212, 214–218, 221, 223–226, 228, 230, 239, 243, 245, 247–249, 271, 274, 276, 352, 354, 373, 375, 401
- Separation assumption, 3, 4, 10, 11
- Shared understanding, 24, 42, 183, 184, 193, 194, 204, 205, 491
- Simple logistic sum function, 95, 131
- Simulated annealing, 394, 412, 413, 415
- Simulation experiment, 22–24, 26, 27, 38, 59, 99, 106, 131, 133, 138, 150, 154, 168, 173, 260, 286, 298, 323, 324
- Situation characteristics, 395–397, 400

- Small world network, 286, 288, 289, 317
 - Social, 3, 4, 8, 9, 11, 13–15, 18, 19, 22, 24, 25, 37
 - Social contagion, 9, 41, 66, 106, 107, 119, 120, 184, 185, 188, 192, 286
 - Social interaction, 9, 15, 18, 22, 26, 27, 36, 39–42, 66
 - Socially aware, 28, 385, 464–466, 470
 - Social media, 15, 22, 26, 40, 184, 200, 204, 300, 301, 312
 - Social neuroscience, 23–25, 37, 99, 119, 184, 203, 263, 265, 491
 - Social response, 25, 197, 205, 236, 243, 263, 264, 383, 385
 - Societal developments, 27, 28
 - Societal perspective, 27
 - Socio-technical system, 22, 28, 491
 - Somatic marker hypothesis, 115, 160
 - Sort, 17, 108, 364, 365, 439
 - Speed factor, 19, 21, 22, 39, 46, 49, 56, 61, 82, 90, 129, 330, 357, 402
 - Speed property, 379
 - Standard combination function, 20, 36, 39, 46, 48, 49, 55, 59, 69, 77
 - State, 5–7, 9, 10, 12, 13, 18, 19, 21, 25
 - State comparison property, 380
 - State-determined system, 16, 17, 22, 37, 39, 46, 85, 86, 88–90, 99, 423
 - State-determined system assumption, 85, 439, 446
 - State relation, 358, 360, 363, 364, 366, 368, 369, 371, 372
 - State relation property, 375
 - Stationary point, 323–328, 337, 338, 340, 345, 346, 375, 377, 378, 379
 - Steepness, 71, 73–75, 83, 92, 131, 133, 146, 151, 168, 227, 260, 396, 400
 - Stochastic world, 159, 166, 171, 176
 - Strong tie, 292–294, 297
 - Strong tie hypothesis, 294
 - Sum function, 20, 49, 74, 75, 82, 92, 95, 216, 247, 311, 346
 - Support model, 470
 - Suppressing, 23–25, 111, 128, 131, 135
 - Symmetric, 60, 81, 91, 95, 96, 251
 - Systematic transformation, 39, 49, 58
- T**
- Taylor approximation, 423, 449, 460
 - Temporal, 4, 15, 19, 23, 25, 27, 36, 39, 46, 49, 59, 89, 92, 99, 109, 114
 - Temporal-causal network, 4, 18–21, 23–28, 36–39, 42, 47
 - Temporal separation, 4, 10, 11
 - Temporal trace language TTL, 387
 - Tendency to motion, 433
 - Theory of Mind (ToM), 195, 465
 - Things of the soul, 432, 435
 - Threshold, 71, 73, 84, 92, 131, 133, 151, 165, 227, 252, 255, 256, 260, 262, 306, 345, 357, 396, 403
 - Time representation, 365
 - Timing of causal effects, 36
 - Trace comparison property, 382, 383
 - Trace selection property, 383
 - Transition system, 28, 421, 423, 439–441, 443, 460, 489
- U**
- Ubiquitous computing, 28
 - Universal quantifier, 377, 387, 388
 - Upward bounded, 79, 80
- V**
- Validation, 38, 43, 138, 229, 256, 263, 347, 350, 355, 356, 358, 479
 - Valuing, 8, 24, 26, 107, 118, 157–160, 166, 167, 177, 179, 198, 271, 282
 - Verification, 26, 38, 39, 252, 323, 324, 327, 329, 347, 350, 351, 390
 - Vetoing, 25, 210, 213, 220, 222, 227, 230
 - Vis viva, 434
- W**
- Weak tie, 292, 293, 297
 - Weak tie hypothesis, 293, 298
 - Winner takes it all competition, 132, 138, 147
 - World state, 17, 51, 56, 81, 164, 166, 167, 171, 193, 195, 213, 216, 273, 276, 344, 374, 375, 401, 443, 456, 457