

Cognitive Systems Monographs

Volume 33

Series editors

Rüdiger Dillmann, University of Karlsruhe, Karlsruhe, Germany
e-mail: ruediger.dillmann@kit.edu

Yoshihiko Nakamura, Tokyo University, Tokyo, Japan
e-mail: nakamura@ynl.t.u-tokyo.ac.jp

Stefan Schaal, University of Southern California, Los Angeles, USA
e-mail: sschaal@usc.edu

David Vernon, University of Skövde, Skövde, Sweden
e-mail: david@vernon.eu

About this Series

The Cognitive Systems Monographs (COSMOS) publish new developments and advances in the fields of cognitive systems research, rapidly and informally but with a high quality. The intent is to bridge cognitive brain science and biology with engineering disciplines. It covers all the technical contents, applications, and multidisciplinary aspects of cognitive systems, such as Bionics, System Analysis, System Modelling, System Design, Human Motion, Understanding, Human Activity Understanding, Man-Machine Interaction, Smart and Cognitive Environments, Human and Computer Vision, Neuroinformatics, Humanoids, Biologically motivated systems and artefacts Autonomous Systems, Linguistics, Sports Engineering, Computational Intelligence, Biosignal Processing, or Cognitive Materials as well as the methodologies behind them. Within the scope of the series are monographs, lecture notes, selected contributions from specialized conferences and workshops.

Advisory Board

Heinrich H. Bülthoff, MPI for Biological Cybernetics, Tübingen, Germany

Masayuki Inaba, The University of Tokyo, Japan

J.A. Scott Kelso, Florida Atlantic University, Boca Raton, FL, USA

Oussama Khatib, Stanford University, CA, USA

Yasuo Kuniyoshi, The University of Tokyo, Japan

Hiroshi G. Okuno, Kyoto University, Japan

Helge Ritter, University of Bielefeld, Germany

Giulio Sandini, University of Genova, Italy

Bruno Siciliano, University of Naples, Italy

Mark Steedman, University of Edinburgh, Scotland

Atsuo Takanishi, Waseda University, Tokyo, Japan

More information about this series at <http://www.springer.com/series/8354>

M. Guadalupe Sánchez-Escribano

Engineering Computational Emotion—A Reference Model for Emotion in Artificial Systems

 Springer

M. Guadalupe Sánchez-Escribano
Escuela Técnica Superior de Ingenieros
Industriales
Universidad Politécnica de Madrid
Madrid
Spain

Advisors:

Ricardo Sanz
Ramón Galán

ISSN 1867-4925 ISSN 1867-4933 (electronic)
Cognitive Systems Monographs
ISBN 978-3-319-59429-3 ISBN 978-3-319-59430-9 (eBook)
DOI 10.1007/978-3-319-59430-9

Library of Congress Control Number: 2017941487

© Springer International Publishing AG 2018

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. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

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

*Under the consent of my beloved David who has loved me and has supported all my hours of study. And the consent of my parents, my sister and my family, that has given me all their love trying to understand me...
...let me to dedicate this work to my beloved grandparents and grandmothers, who might be taking care of me from the sky.
I miss you.
To: Amalia, Agustin, Nati and Eloy*

Acknowledgements

I would like to thank Ricardo Sanz for the opportunity he gave me to discover this amazing field. Without his willingness to share knowledge, resources, histories, and experience, this research would not have ever been possible. He has been an excellent friend, mentor, and teacher.

In the same vein, I would like to extend great thanks to Ramón Galán who always offered me support and comprehension in my silence. I started within the research field thanks to him, and I have to thank him everything I learnt from his expertise. My first steps in code and systems were built from the basis of working with him. He has been the perfect second advisor and mentor.

I hope to continue my research with them.

Throughout this time, one person has always been there during those difficult and trying times. She has pushed toward improving my work and offering support and help to me. I would like to express gratitude toward Paloma de la Puente, the best engineer, colleague, and friend you might find.

This text shows a piece of research that can be read under the correct understanding because of the input, influence, text improvement, and expert knowledge of Eva R. Porras. I will owe a great deal of gratitude towards you.

Most importantly of all, I show extensive gratitude to all of the people who contributed to this moment in my life. Great thanks to Jose Luis García Vidueira and Gonzalo Sánchez Casado.

Great thanks to all

Contents

1	Introduction	1
1.1	Background	2
1.2	Rationale and Hypotheses	3
1.3	Framework	4
1.4	Questions and Objectives	5
1.5	Dissertation Structure	8
1.6	Assumptions, Limitations, and Scope	10
2	Models	11
2.1	Introduction	12
2.2	Agents	12
2.3	The Emotional Agent	16
2.4	Technological Portrait	20
2.5	On Appraisal	22
2.6	The Exploitation of Models	24
2.7	Self-Adaptive Systems	26
2.8	Conclusions	29
3	State of the Art	31
3.1	Introduction	31
3.2	General Overview	32
3.3	Trends in Biological Emotion Research	36
3.3.1	Research on the Neural Bases of Emotion	36
3.3.2	Psychology and Emotion Research	38
3.4	Trends in Artificial Emotion Research	40
3.4.1	Emotion in Artificial Intelligence and Robotics	41
3.4.2	Emotion and Cognitive Science	44
3.5	Emotion Modeling	50
3.6	Discussion	56

4	Principles for Computational Emotion	59
4.1	Introduction	60
4.2	The State of the Challenge.	62
4.3	Conceptualization of Artificial Emotion	64
	4.3.1 The Cognitive Approach of Emotion	65
	4.3.2 Emotion and Abstractions.	68
	4.3.3 Computational-Emotion and Computational-Feeling	70
4.4	Appraisal and Computational Emotion.	72
	4.4.1 Computational Appraisal	75
	4.4.2 Computational Process of Appraisal.	77
	4.4.3 Measurement of Compliance	82
4.5	Perception	86
4.6	Autobiographical-Self.	90
	4.6.1 The Autobiographical Knowledge	91
	4.6.2 Relationships Between Memories and Self.	92
	4.6.3 Consequences of Relationships Between Memories and Self	94
	4.6.4 Self-Memory System (SMS).	94
	4.6.5 Grounds of the Self	96
4.7	Representation and Computational Feeling	100
	4.7.1 The Inner-Environment.	104
	4.7.2 Grounds for Emotion and Feeling	106
4.8	Architectural Theoretical Foundations	114
4.9	Alertness, Arousal and Action Readiness.	119
4.10	Discussion	120
5	A Model of Computational Emotion	121
5.1	Introduction	121
5.2	Previous Works	125
5.3	Thesis	127
5.4	Emotional Agent AGSys	129
	5.4.1 Conceptualization of ESys	132
	5.4.2 External Environment.	134
	5.4.3 Inner Environment	137
	5.4.4 Ongoing Interaction Between Environments.	142
5.5	Principles of Perception in AGSys.	143
	5.5.1 The Inner-Object	144
	5.5.2 The Process of Perception	146
	5.5.3 Principles for Emotion Perception	149
	5.5.4 Principles and Baselines for Representation	151
	5.5.5 Summarized Problem	156
5.6	The Inner Agent ESys	156
	5.6.1 The Appraisal.	158
	5.6.2 The Emotional Goals	164
	5.6.3 Valenced Reactions	179

- 5.7 Feeling in AGSys. 180
 - 5.7.1 The Problem of Representation in AGSys 183
 - 5.7.2 Regarding the Problem of Autonomy in AGSys. 187
 - 5.7.3 The R-Loop 190
 - 5.7.4 MEP Model 193
- 5.8 Architectural Foundations 207
 - 5.8.1 The Environments of AGSys 207
 - 5.8.2 AGSys–ESys Reciprocal Actions 209
 - 5.8.3 Essential Questions Regarding the Appraisal 210
 - 5.8.4 Questions About Reconfiguration 212
 - 5.8.5 About Agency 216
 - 5.8.6 Generalization. 216
- 5.9 Conclusions 217
- 6 Discussion 221**
 - 6.1 Introduction 221
 - 6.2 The Approach 223
 - 6.3 Emotion Based Autonomy 225
 - 6.4 Adaptive Systems. 227
 - 6.5 What Is Artificial Emotion. 231
 - 6.6 Conclusion. 233
- 7 Future Works 235**
 - 7.1 The Approach 236
 - 7.2 Emotion Research 237
- Glossary 241**
- Declaration 247**
- References 249**

List of Figures

Fig. 2.1	Minimal characteristics to derive four types of agents according to Nwana [187]	13
Fig. 2.2	Classification of Software Agents according to Nwana [187] . . .	14
Fig. 2.3	Conceptualization of an artificial agent (adaptation from the viewpoint of Russell and Norvig [222]).	15
Fig. 2.4	Structure of an Emotional cognitive agent under the approach proposed in this thesis.	18
Fig. 2.5	Autonomic control loop. (<i>Source</i> Brun et al. [37]).	19
Fig. 2.6	The vision of self-adaptive software under Vogel and Giese [262]. (<i>Source of figure</i> Vogel and Giese [262]).	28
Fig. 2.7	Runtime models perspective by Vogel and Giese [262]. (<i>Source of figure</i> Vogel and Giese [262]).	28
Fig. 2.8	Layered architecture for an adaptation engine by Vogel and Giese [262]. (<i>Source of figure</i> Vogel and Giese [262])	29
Fig. 3.1	Kismet is an autonomous robot developed by Breazeal [32] at MIT Artificial Intelligence Laboratory: a Kismet displays an angry expression, b Kismet displays surprise, c Framework for building the motivational system. <i>Source of figures</i> Breazeal [32].	42
Fig. 3.2	Cog is an upper-torso humanoid robot: a cog is a platform with twenty-one degrees of freedom and sensors enough to approach human movements and senses, b range of motion for the neck and torso. <i>Source of figures</i> Brooks et al. [36]	43
Fig. 3.3	The sequence of “Wamoeba” robots from left to right: Wamoeba-2, Wamoeba-2R, Wamoeba-2Ri and Wamoeba-3 . . .	43
Fig. 3.4	FUWA (Fundan Intelligent Robot service developed at the Fudan University in Shanghai (China): a the robot shaking hands with human, b the frame of emotion generation, where event perceptions are sent to an evaluation model, in which the state of robot’s body and goals are taking into	

account at the same time, and *c* inspired in OCC model by Ortony et al. [198], the extern stimuli are classified into several events according to some defined features.

Source of figures Qingji et al. [212] 44

Fig. 3.5 General view of SOAR cognitive architecture 46

Fig. 3.6 General view of LIDA cognitive architecture 46

Fig. 3.7 The ACT-R 5.0 model. *Source* [46] 47

Fig. 3.8 Subsumption Architecture by Brooks et al. [35], used to implement physical robots and purely reactive software agents. The architecture is made up of modules placed in layers, together with augmented finite state machines. *Source* Nwana [187]. 47

Fig. 3.9 4D/RCS reference model architecture: Real-time control system (RCS) is a cognitive architecture designed to enable any level of intelligent behavior. The first version of RCS was developed as an sensory-interactive goal-directed controller. It has evolved into 4D/RCS over the years, an intelligent controller for industrial robots and several other intelligent tools 48

Fig. 3.10 General overview of SHAME architecture: an event appraiser compute an emotion impulse vector (EIV) which is normalized to an emotion vector (NEIV) that acts as an impulser. *Source* [208] 50

Fig. 3.11 General overview of MicroPSI architecture 50

Fig. 3.12 General overview of a computational instantiation of the cognitive-motivational-emotive system for EMA made by Gratch and Marsella [94]. EMA was named in honor of the book of [242] which was tilted ‘Emotion and Adaptation’. This work proposed a unified view of appraisal and coping. *Source* Gratch and Marsella [94] 53

Fig. 3.13 General overview of TABASCO: the goal is the integration of the emotion process within the architecture of an agent, and model emotions as adaptive processes related to the agent-environment interaction. TABASCO’s essential idea is that conceptual processing is not the only responsible of sensory-motor, schematic and conceptual processing. It is also affected by the generation of action, as it was proposed by Leventhal [146] in his theory ‘perceptual motor theory of emotion.’ *Source* Staller and Petta [245] 53

Fig. 3.14 Model of Emotions of Frijda [80] under the central notion of concerns. The two architectures artificial concern realisation system (ACRES) and the later refinement

	named Acres WILL, began as an implementation of Frijda [80] theory of emotions	54
Fig. 3.15	ALMA is part of the VirtualHuman Project to develop interactive characters with conversational skills. ALMA provides a personality profile, real-time emotions and moods: a screenshot of the VirtualHuman system, b pull and push mood change function (<i>source of figures a</i> and b Gebhard [87]), and c Mood changes in ALMA. <i>Source of figure:</i> https://www.dfki.de	54
Fig. 3.16	Global structure of emotion types in OCC by Ortony et al. [198].	55
Fig. 4.1	Comparative overview of major appraisal dimensions as postulated by different theorists from the source of Davidson et al. [61]	73
Fig. 4.2	Environmental correlation and cognitive equivalence under the vision of Lòpez Paniagua [151]	87
Fig. 4.3	The Autobiographical memory knowledge-base (<i>Source</i> Conway and Pleydell-Pearce [50]).	91
Fig. 4.4	Knowledge structures in autobiographical model (<i>Source</i> Conway [49])	99
Fig. 5.1	Tools for explanation are required in order to build correspondences between theories and approaches	123
Fig. 5.2	General view of a previous conceptualization regarding the ideas of this thesis	126
Fig. 5.3	Thesis.	127
Fig. 5.4	Thesis.	128
Fig. 5.5	The abstract model of an open and a closed system from the point of view of Wang (adaptation from Wang [266])	131
Fig. 5.6	Essentials and main structures of the emotional AGSys according to the approach proposed in this work, in order to perform computational emotion	133
Fig. 5.7	Conceptualization of Artificial Emotion and Computational Emotion (see rationale ♣ 7)	133
Fig. 5.8	Conceptualization of environment reconfiguration: Θ_i represents the normal work of AGSys, and Θ_j the one related to an environment that conveys the situational state of fear to AGSys	135
Fig. 5.9	Illustration of a possible AGSys context model for action readiness, on the basis of the MLContext’s abstract syntax metamodel of Hoyos et al. [111]	138
Fig. 5.10	Structure of the emotional AGSys that shows the integration of the Error-Reflex agents	140

Fig. 5.11 A generic program for an Error-Reflex-Agent (adaptation made on the reflex agent program by Russell and Norvig [222]). 140

Fig. 5.12 Ongoing interaction between contexts 141

Fig. 5.13 Inner environment conceptualization. 142

Fig. 5.14 Conceptualization of an inner-Object Representation: **a** From the point of view of AGSys, and **b** according to an easier conceptual representation. 145

Fig. 5.15 Conceptualization about perception recursiveness and inner-Object growth (I). 147

Fig. 5.16 Conceptualization about perception recursiveness and inner-Object growth (II). 148

Fig. 5.17 Conceptualization of Computational Emotion (c-Emotion) and Computational Feeling (c-Feeling). 151

Fig. 5.18 The essential conceptualizations of an external object inside AGSys 152

Fig. 5.19 Partial Model of an external object Structure inside AGSys 153

Fig. 5.20 Analogies with the terminology of Damasio [54], focusing on defining comparable computational artifacts that perform similar functions in artificial systems 154

Fig. 5.21 Two concerns cannot be concurrently appraised. However, there can be several parallel appraisals for the same concern 159

Fig. 5.22 Conceptualization of a-Dimension, c-Concern and c-emoObject. 161

Fig. 5.23 A single conceptualization of a generic *Universe of Emotion* described under dimensions of assessment. 162

Fig. 5.24 A single view of an agent based goal (*Source* under the common understanding of typical agent based on Macal and North [154]). 166

Fig. 5.25 A single view of the agent based structure of an emotional goal. 167

Fig. 5.26 An extended view of the agent based structure of an emotional goal and the relationship with the *Emotion Attribute* vectors. 168

Fig. 5.27 An extended view of the agent based structure of an emotional goal and the relationship with the *Emotion Attribute* vectors. 169

Fig. 5.28 Structure of the system of emotional goals within an equivalent structure: **a** illustrates a view according to Structured Systems Analysis and Design Methodology (SSADM), and **b** illustrates a view according to SysML in order to clarify generalizations 170

Fig. 5.29 Conceptualization of error cause-effect. 171

Fig. 5.30 Conceptualization of an emotional error damper system. 172

Fig. 5.31 Quality of c-Value 178

Fig. 5.32 Conceptualization of a state of action readiness query, which comes from a concrete c-emoGoal. 180

Fig. 5.33 Conceptualization made in this thesis regarding the threefold representation of an external object: *inner-Object*, *emotion-Object* and *feeling-Object* 182

Fig. 5.34 Generic taxonomy of the Self-Memory-System and its main concepts and properties according to the theory of Conway and Pleydell-Pearce 186

Fig. 5.35 Taxonomy of a generic Autobiographical Memory Knowledge Base: **a** Classification of concepts according to the theory of Conway [49] (Fig. 4.4 shows the conceptual view of the authors), **b** Conceptualization of a General Event structure, and **c** Conceptualization of an Episodic Memory structure 188

Fig. 5.36 Conceptualization of the *Self-Memory-System* operation by using the work of Vogel and Giese 190

Fig. 5.37 Conceptualization of the three assumed artifacts in order to approach the experience of emotion 192

Fig. 5.38 Conceptualization of an operative unit in *R-Loop*. 193

Fig. 5.39 MEP builds the feeling-Object by means of three recursions of R-Loop: **a** The Physical-based process, **b** cp-Emotion process, and **c** cp-Feeling process 196

Fig. 5.40 Conceptualization of the first stage in MEP 197

Fig. 5.41 Conceptual model of Basic regulation as a graph of instances. 198

Fig. 5.42 Conceptualization of the cp-Emotion stage in MEP 200

Fig. 5.43 Conceptualization of the cp-Feeling stage in MEP 201

Fig. 5.44 Conceptualization of the representational stage in MEP 204

Fig. 5.45 General view of the proposed solution aimed to approach the emotion representation 206

Fig. 5.46 General view of two representations made on the basis of the same external object. 208

Fig. 5.47 General view of AGSys–ESys interactions. 209

Fig. 5.48 Overview related to the conceptualization of the three proposed environments. 210

Fig. 5.49 General view of the structure of and the reciprocal interaction between AGSys and ESys. AGSys logic has been included in order to illustrate the interaction with the context reconfiguration and the memory system. 211

Fig. 5.50 Once MEP starts the second recursion related to the cp-Emotion process, the concurrency of processes related to additional objects is not feasible. 212

Fig. 5.51 Figure **a** shows a general view of the appraisal work related to a concrete object. Figure **b** describes a possible essential structure of the intensity detector that obtains the higher values of x_1 and x_2 . And figure **c** describes the interfaces between these two detectors showing a connector that joins this property via standard ports. 213

Fig. 5.52 General overview of the reconfiguration strategies in *AGSys* and *ESys* 214

Fig. 5.53 General application of the strategy *AGSys-ESys* in a layered architecture. 218

Abstract

Emotion is generally argued to be an influence on the behavior of life systems, largely concerning flexibility and adaptivity. The way in which life systems act in response to particular situations of the environment has revealed the decisive and crucial importance of this feature in the success of behaviors. And this becomes a source of inspiration for artificial systems design.

During the last decades, artificial systems have undergone such an evolution that each day more are integrated in our daily life. The subsequent effects are related to an increased demand of systems that ensure resilience, robustness, availability, security, or safety among others. These attributes raise fundamental challenges in control and complex systems design.

This thesis has been developed under the framework of the Autonomous System project, a.k.a the *ASys Project*. Short-term objectives of immediate application focus on the design of improved systems, seeking more intelligence in control strategies. Besides this, long-term objectives underlying *ASys Project* concentrate on high-order capabilities such as cognition, awareness, and autonomy.

This work is placed within the general fields of *engineering* and *emotion science* and provides a theoretical foundation for engineering and designing computational emotion for artificial systems. The starting research question that has grounded this thesis aims at addressing the problem of emotion-based autonomy. And how to feedback systems with valuable meaning conforms the general objective. Both the starting question and the general objective have underlaid the study of emotion, the influence on systems behavior, the key foundations that justify this feature in life systems, how emotion is integrated within the normal operation, and how this entire problem of emotion can be explained in artificial systems. By assuming essential differences concerning structure, purpose, and operation between life and artificial systems, the essential motivation has been the exploration of what emotion solves in nature to afterward analyze analogies for man-made systems.

The approach provides a reference model in which a collection of entities, relationships, models, functions, and informational artifacts are all interacting to provide the system with non-explicit knowledge under the form of emotion-like relevances. This solution aims to provide a reference model as a framework to

design solutions for emotional operation, but related to the real needs of artificial systems.

The proposal consists of a multi-purpose architecture that implements two broad modules in order to attend: (a) the range of processes related to the environment affectation, and (b) the range or processes related to the emotion perception-like and the higher levels of reasoning. This has required an intense and critical analysis beyond the state of the art around the most relevant theories of emotion and technical systems, in order to obtain the required support for those foundations that sustain each model.

The problem has been interpreted and is described on the basis of *AGSys*, an agent assumed with the minimum rationality as to provide the capability to perform emotional assessment. *AGSys* is a conceptualization of a *model-based cognitive agent* that embodies an inner agent *ESys*, the responsibility for performing the emotional operation inside of *ASys*.

The solution consists of multiple computational modules working federated and aimed at conforming a mutual feedback loop between *AGSys* and *ESys*. Throughout this solution, the environment and the effects that might have an influence on the system are described as different problems. While *AGSys* operates as a common system within the external environment, *ESys* is designed to operate within an *inner environment* built on the basis of those relevances that might occur inside of *AGSys*. This allows for a high-quality separated reasoning concerning mission goals defined in *AGSys* and emotional goals defined in *ESys*. This way, a possible path for high-level reasoning under the influence of goals congruence is provided.

The high-level reasoning model uses knowledge about emotional goals stability, hence allowing for new directions in which mission goals might be assessed according to the situational state of this stability. This high-level reasoning is grounded by the work of *MEP*, a model of emotion perception that is conceived as an analogy of a well-known theory in emotion science. The work of this model is described under the operation of a recursive-like process denoted *R-Loop*, together with a system of emotional goals that are assumed to be individual agents. This way, *AGSys* integrates knowledge about the relation between a perceived object and the effect which this perception induces on the situational state of the emotional goals. This knowledge might afford a new system of information that might provide the sustainability for a high-level reasoning. The extent to which this reasoning might be approached is just delineated and assumed as future work.

This thesis has been studied from a wide variety of fields of knowledge. These knowledge categories can be classified into two main area groups: (a) the fields of psychology, cognitive science, neurology, and biological sciences in order to obtain understanding concerning the problem of the emotional phenomena, and (b) a large amount of computer science branches such as autonomic computing (AC), self-adaptive software, self-X systems, model-integrated computing (MIC), or the paradigm of Model Driven Engineering among others, in order to obtain knowledge about tools for designing each part of the solution.

The final approach has been mainly performed on the basis of the entire acquired knowledge and described within the fields of artificial intelligence, Model-Based Systems (MBS), and additional mathematical formalizations to provide punctual understanding in those cases that have required it. This approach describes a reference model to feedback systems with valuable meaning, allowing for reasoning with regard to (a) the relationship between the environment and the relevance of the effects on the system, and (b) dynamical evaluations concerning the inner situational state of the system as a result of those effects.