

# Agents with Affective Traits for Decision-Making in Complex Environments



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

Author: Bexy Alfonso Espinosa

Supervisors: Dr. Emilio P. Vivancos Rubio  
Dr. Vicente J. Botti Navarro

A thesis submitted for the degree of  
*Doctor por la Universitat Politècnica de València*

Departamento de Sistemas Informáticos y Computación  
Universitat Politècnica de València  
Valencia, Spain  
September 2017



# PhD Thesis

**Title:** Agents with Affective Traits for Decision-Making in Complex Environments

**Author:** Bexy Alfonso Espinosa

**Supervisors:** Dr. Emilio P. Vivancos Rubio  
Dr. Vicente J. Botti Navarro

**Reviewers:**  
Dr. Garazi Azanza  
Dr. Paolo Novais  
Dr. Viviana Patti

**Examination board:**

**President:** Dr. Vicente J. Inglada

**Secretary:** Dr. Miguel Arevalillo

**Member:** Dr. Viviana Patti

**Day of the defense:** September 21<sup>th</sup>, 2017



*A mi familia*



---

## Acknowledgments

Es casi imposible realizar un proyecto a largo plazo sin apoyo alguno y terminar ileso. Éste no está exento. Muchas palabras de aliento y mucho apoyo en todos los sentidos han estado detrás de este trabajo.

Quiero destacar, ante todo, el apoyo de mis directores Emilio y Vicente, quienes, además de ser unos profesionales e investigadores excelentes, también, y ante todo, son excelentes personas. A Vicente quiero agradecerle por brindarme la oportunidad de trabajar en la investigación como un miembro más del GTI, lo cual ha sido una experiencia única y muy enriquecedora para mi. A Emilio por su apoyo, comprensión e implicación en todos los sentidos en los momentos necesarios, y por estar siempre a mano de una reunión improvisada o de un Skype. También quiero agradecer a Ximo, con quien hemos conformado un gran equipo y de quien también he tenido ayuda y palabras de aliento en los momentos más críticos.

I would also like to thank Stacy Marcella for allowing me to work with him and with Margot Lhommet on the CESAR lab at the Northeastern University. I am deeply grateful for the time they devoted to me and for their collaboration together with David Pynadath. It was a really enriching experience in every sense. On his research group I found, not only really good researchers, but also great human beings. Thanks to Stacy, Sophie, Yuyu, Dan, and Margot for their warm reception.

El llevar a término este trabajo también se debe al buen ambiente de trabajo del que he tenido la suerte de disfrutar. Por ello me gustaría agradecer a todos los miembros de GTI. Al trabajar con ellos se tiene todo en un pack: compañerismo, amabilidad, talento y humanismo. Especialmente quisiera

---

agradecer a mis antiguos compañeros de laboratorio Mario, Ricard, Raúl y Juan Andrés, con quienes he tenido la suerte de compartir un ambiente de trabajo excelente y además una buena amistad. También agradecer a mis compañeros Elena, Stella, Joan, Juanmi, María, Jaime, Jaume, Javi, Mohanad, Ângelo y a todos los que en algún momento tuvieron unas palabras de aliento y apoyo para no desfallecer antes de llegar al final del camino.

Por ultimo, y por ello no menos importante, quisiera agradecer a mi familia y resto de amigos, que aunque algunos están en la distancia, siempre han sido mi fortaleza y mi fuente de inspiración. Especialmente a mi madre, quien consciente o inconscientemente, siempre ha sido un motor y un ejemplo de espíritu de lucha contra todo tipo de adversidad y contra todo pronóstico negativo. Gracias también a Mari Carmen, Paco, Águeda, Derik, quienes estuvieron ahí para echar una mano o dar ánimos cuando se necesitaba. Y gracias a David, quien también es autor indirecto de este trabajo por el ánimo y el apoyo que me ha dado y sin el cual el final de este camino hubiese sido más incierto.



---

---

## Summary

Recent events have probably lead us to wonder why people make decisions that seem to be irrational, and that go against any easily understandable logic. The fact that these decisions are emotionally driven often explains what, at first glance, does not have a plausible explanation. Evidence has been found that proves that emotions and other affective characteristics guide decisions beyond a purely rational deliberation. Understanding the way emotions take place, the way emotions change, and/or the way emotions influence behavior, has traditionally been a concern of several fields including psychology and neurology. Moreover, other sciences such as behavioral economics, artificial intelligence, and in general, all sciences that aim to understand, explain, or simulate human behavior, acknowledge the important role of affective characteristics in this task.

Specifically, artificial intelligence uses psychological findings in order to create agents that simulate human behavior. Nevertheless, individual research efforts in modeling affective characteristics are often overlapped, short of integration, and they lack of a common conceptual system. This deprives individual researches of the exchange and cooperation's inherent benefits, and makes the task of computationally simulating affective characteristics more difficult. Although much individual effort has been put in classifying, formalizing and modeling emotions and emotion theories on some fields, recognized researchers of emotions' and affective processes' modeling report that a common formal language, an informal conceptual system, and a general purpose affective agent architecture will greatly improve the interdisciplinary exchange and the intradisciplinary coordination.

---

The research literature proposes a wide amount of affective models that deal with some of: relationship between emotions and cognition, relationship between emotions and behavior, emotions and their evolutionary account, emotions for appraising situations, emotion regulation, etc. These models are useful tools for addressing particular emotion-related issues. Furthermore, computational approaches that are based on particular psychological theories have also been proposed. They often address domain specific issues starting from a specific psychological theory. In such solutions, the absence of a common conceptual system and/or platform, makes difficult the feedback between psychological theories and computational approaches.

This thesis systematizes and formalizes affect-related theories, what can benefit the interdisciplinary exchange, the intradisciplinary coordination, and hence, allows the improvement of involved disciplines. Specifically this thesis makes the following contributions: (1) a theoretical framework that includes the main processes and concepts that a model of an affective agent with practical reasoning should have; (2) a general-purpose affective agent architecture that shares the concepts of the proposed theoretical framework; (3) an implementation-independent formal language for designing affective agents that have the proposed architecture; and (4) a specific agent language for implementing affective agents which is an extension of a BDI language.

Some studies with human participants have helped to validate the contributions of this thesis. They include classical games of game theory, and a study with 300 participants, which have provided the necessary information to evaluate the contributions. The validation has been performed in three directions: determine whether the proposed computational approach represents better the human behavior than traditional computational approaches; determine whether this approach allows to improve psychological theories used by default; and determine whether the proposed affective agents' behavior is closer to human behavior than the behavior of a purely rational agent.

---

---

## Resumen

Probablemente algunos eventos recientes nos han conducido a preguntarnos por qué las personas toman decisiones aparentemente irracionales y en contra de alguna lógica fácilmente comprensible. El hecho de que estas decisiones estén bajo la influencia de las emociones a menudo explica lo que, a primera vista, parece no tener una explicación aceptable. En este sentido, se han encontrado evidencias que prueban que las emociones y otras características afectivas condicionan las decisiones más allá de una deliberación meramente racional. Entender cómo las emociones tienen lugar, cómo cambian y cómo influyen en el comportamiento, ha sido tradicionalmente de interés para muchos campos de investigación, incluyendo la psicología y la neurología. Además, otras ciencias como la economía conductual o la inteligencia artificial reconocen el importante papel de las características afectivas en esta tarea.

Específicamente, la inteligencia artificial utiliza los resultados obtenidos en psicología para crear agentes que simulan el comportamiento humano. Sin embargo, a menudo los esfuerzos individuales de investigación en el modelado del afecto se solapan, carecen de la suficiente integración y de un sistema conceptual común. Esto limita a las investigaciones individuales para disponer de los beneficios que ofrecen el intercambio y la cooperación, y hace más compleja la tarea de simular los procesos afectivos. Las emociones y teorías relacionadas han sido clasificadas, formalizadas y modeladas. No obstante, reconocidos investigadores argumentan que un lenguaje formal común, un sistema conceptual informal y una arquitectura de agentes de propósito general, mejorarán significativamente el intercambio interdisciplinar y la coordinación intradisciplinar.

---

En la literatura se propone una amplia cantidad de modelos afectivos que modelan: la relación entre las emociones y la cognición, la relación entre las emociones y el comportamiento, las emociones para evaluar las situaciones, la regulación de emociones, etc. Estos modelos son herramientas útiles para abordar aspectos particulares relacionados con las emociones. Además, se han realizado propuestas computacionales que abordan aspectos específicos sobre la base de teorías psicológicas específicas. En éstas soluciones, la ausencia de una plataforma y/o sistema conceptual dificulta la retroalimentación entre las teorías psicológicas y las propuestas computacionales.

Esta tesis sistematiza y formaliza teorías relacionadas con el afecto, lo cual beneficia el intercambio interdisciplinar y la coordinación intradisciplinar, y por tanto, permite el desarrollo de las disciplinas correspondientes. Específicamente esta tesis realiza las siguientes contribuciones: (1) una plataforma teórica que incluye los conceptos y procesos principales que debería poseer un modelo de agentes afectivos con razonamiento práctico; (2) una arquitectura de agentes de propósito general que comparte los conceptos de la plataforma teórica propuesta; (3) un lenguaje formal independiente de la implementación, para diseñar agentes afectivos que poseen la arquitectura propuesta; y (4) un lenguaje de agentes específico para implementar agentes afectivos el cual es un extensión de un lenguaje BDI.

Algunos estudios con participantes humanos han ayudado a validar las contribuciones de esta tesis. Estos incluyen juegos clásicos de teoría de juegos y un estudio con 300 participantes, los cuales han proporcionado la información necesaria para evaluar las contribuciones. La validación se ha realizado en tres direcciones: determinar si la propuesta computacional que se ha realizado representa mejor el comportamiento humano que propuestas computacionales tradicionales; determinar si esta propuesta permite mejorar las teorías psicológicas empleadas por defecto; y determinar si el comportamiento de los agentes afectivos propuestos se acerca más al comportamiento humano que el comportamiento de agentes racionales.

---

---

## Resum

Probablement alguns esdeveniments recents ens han conduït a preguntar-nos per què les persones prenen decisions que aparentment són irracionals i que van en contra d'algun tipus de lògica fàcilment comprensible. El fet que aquestes decisions estiguin sota la influència de les emocions sovint explica el que, a primera vista, sembla no tenir una explicació acceptable. En aquest sentit, s'han trobat evidències que provenen que les emocions i altres característiques afectives condicionen les decisions més enllà d'una deliberació merament racional. Entendre com les emocions tenen lloc, com canvien i com influeixen en el comportament, ha estat tradicionalment d'interès per a molts camps d'investigació, incloent la psicologia i la neurologia. A més, altres ciències com l'economia conductual, la intel·ligència artificial i, en general, totes les ciències que intenten entendre, explicar o simular el comportament humà, reconeixen l'important paper de les característiques afectives en aquesta tasca.

Específicament, la intel·ligència artificial utilitza els resultats obtinguts en psicologia per crear agents que simulen el comportament humà. No obstant això, sovint els esforços individuals d'investigació en el modelatge de l'afecte es solapen, no tenen la suficient integració ni compten amb un sistema conceptual comú. Això limita a les investigacions individuals, que no poden disposar dels beneficis que ofereixen l'intercanvi i la cooperació, i fa més complexa la tasca de simular els processos afectius. Les emocions i teories relacionades han estat classificades, formalitzades i modelades. No obstant això reconeguts investigadors argumenten que un llenguatge formal comú, un sistema conceptual informal i una arquitectura d'agents de propòsit general, milloraran significativament l'intercanvi interdisciplinar i la coordinació intradisciplinar.

---

En la literatura es proposa una àmplia quantitat de models afectius que modelen: la relació entre les emocions i la cognició, la relació entre les emocions i el comportament, les emocions per avaluar les situacions, la regulació d'emocions, etc. Aquests models són eines útils per abordar aspectes particulars relacionats amb les emocions. A més, s'han realitzat propostes computacionals que aborden aspectes específics sobre la base de teories psicològiques específiques. En aquestes solucions, l'absència d'una plataforma i/o sistema conceptual dificulta la retroalimentació entre les teories psicològiques i les propostes computacionals.

Aquesta tesi sistematitza i formalitza teories relacionades amb l'afecte, la qual cosa beneficia l'intercanvi interdisciplinari i la coordinació intradisciplinari, i per tant, permet el desenvolupament de les disciplines corresponents. Específicament aquesta tesi realitza les següents contribucions: (1) una plataforma teòrica que inclou els conceptes i processos principals que hauria de posseir un model d'agents afectius amb raonament pràctic; (2) una arquitectura d'agents de propòsit general que comparteix els conceptes de la plataforma teòrica proposada; (3) un llenguatge formal independent de la implementació, per dissenyar agents afectius que posseeixen l'arquitectura proposada; i (4) un llenguatge d'agents específic per implementar agents afectius el qual és una extensió d'un llenguatge BDI.

Alguns estudis amb participants humans han ajudat a validar les contribucions d'aquesta tesi. Aquests inclouen jocs clàssics de teoria de jocs i un estudi amb 300 participants, els quals han proporcionat la informació necessària per avaluar les contribucions. La validació s'ha realitzat en tres direccions: determinar si la proposta computacional que s'ha realitzat representa millor el comportament humà que propostes computacionals tradicionals; determinar si aquesta proposta permet millorar les teories psicològiques emprades per defecte; i determinar si el comportament dels agents afectius proposats s'acosta més al comportament dels humans que el comportament dels agents racionals.

---

---

# Contents

<b>Contents</b>	<b>xiii</b>
<b>List of Figures</b>	<b>xix</b>
<b>List of Tables</b>	<b>xxiii</b>
<b>I Introduction</b>	<b>1</b>
<b>1 Introduction</b>	<b>3</b>
1.1 Motivation . . . . .	6
1.1.1 Description of the Problematic Situation . . . . .	8
1.2 Hypothesis . . . . .	10
1.3 Goals . . . . .	10
1.3.1 General Goal . . . . .	10
1.3.2 Specific Goals . . . . .	11
1.4 Structure of the Document . . . . .	12
<b>II State of the Art</b>	<b>15</b>
<b>2 State of the Art</b>	<b>17</b>
2.1 Introduction . . . . .	17
2.2 The Influence of Emotions in Cognition and Practical Reasoning	18
2.2.1 Emotions and Memory . . . . .	19
2.2.2 The Influence of Emotions on Behavior . . . . .	20

# CONTENTS

---

2.2.3	Emotions and their Evolutionary Account . . . . .	21
2.2.4	Emotions in the Evaluation of the Situation . . . . .	22
2.2.5	Emotion Regulation . . . . .	24
2.2.6	Evaluative Implications of Emotions . . . . .	26
2.2.7	Emotions and Social Interaction . . . . .	26
2.3	Interindividual Differences . . . . .	27
2.4	Affective Agents with Practical Reasoning . . . . .	30
2.4.1	Conceptual Systems for Intelligent Affective Agents . .	31
2.4.1.1	Cognitive Processes of Intelligent Agents . . . .	33
2.4.1.2	Conceptual Systems . . . . .	34
2.4.2	Use of Core Processes and Attributes in Computational Approaches . . . . .	37
2.4.2.1	Affect and Memory . . . . .	39
2.4.2.2	Affect Influence on Behavior . . . . .	41
2.4.2.3	Emotions and their Evolutionary Account . . . .	44
2.4.2.4	Affect in the Evaluation of the Situation . . . .	47
2.4.2.5	Affect Regulation . . . . .	47
2.4.2.6	Evaluative Implications of Affect . . . . .	49
2.4.2.7	Affect and Social Interaction . . . . .	51
2.5	General Approaches . . . . .	52
2.6	Conclusions . . . . .	54

## III Affective Architecture 59

<b>3</b>	<b>GenIA<sup>3</sup>: a General purpose Intelligent Affective Agent Archi- tecture</b>	<b>61</b>
3.1	Introduction . . . . .	61
3.2	Processes of an Affective Agent Life Cycle . . . . .	62
3.2.1	Processes in a BDI Agent . . . . .	63
3.2.2	Agent's Affective Processes . . . . .	64
3.3	Integration of Affective and Rational BDI processes . . . . .	68
3.4	Agent Attributes for BDI and Affective Processes . . . . .	71



3.4.1	Representation for BDI Processes . . . . .	71
3.4.2	Representation for Affective Processes . . . . .	72
3.4.3	Representation of Affect-Related Attributes in an Agent Language . . . . .	74
3.4.3.1	Affect Generation . . . . .	74
3.4.3.2	Affect Experience . . . . .	76
3.4.3.3	Affect Effects . . . . .	76
3.5	Conclusions . . . . .	77
<b>4</b>	<b>Formalization of GenIA<sup>3</sup></b>	<b>79</b>
4.1	Introduction . . . . .	79
4.2	Extension of the Jason Agent Reasoning Cycle . . . . .	80
4.3	Extension of the Jason Language . . . . .	83
4.4	Extension of the AgentSpeak's Operational Semantics . . . . .	88
4.4.1	New Transition Rules . . . . .	95
4.5	<b>GenIA<sup>3</sup>'s</b> Default Design . . . . .	103
4.6	Conclusions . . . . .	108
<b>IV</b>	<b>Evaluation</b>	<b>111</b>
	<b>Introduction</b>	<b>113</b>
<b>5</b>	<b>Affect and Personality in Classical Games for Addressing So- cial Preferences</b>	<b>115</b>
5.1	Introduction . . . . .	115
5.2	Prisoner's Dilemma . . . . .	116
5.2.1	Rational vs. Human Behavior in the Prisoner's Dilemma	116
5.2.2	<b>GenIA<sup>3</sup></b> Agents' Behavior in the Prisoner's Dilemma .	118
5.3	Trust Game . . . . .	122
5.3.1	Rational vs. Human Behavior in the Trust Game . . . . .	123
5.3.2	<b>GenIA<sup>3</sup></b> Agents' Behavior in the Trust Game . . . . .	123

# CONTENTS

---

<b>6</b>	<b>GenIA<sup>3</sup>'s Agents Playing Black Jack. A Case Study</b>	<b>129</b>
6.1	Black Jack Experiment . . . . .	130
6.1.1	General Description . . . . .	130
6.1.2	Design of the Experiment . . . . .	132
6.1.3	Measures . . . . .	136
6.1.4	Data Analysis . . . . .	137
6.2	Agents Playing Black Jack . . . . .	150
6.2.1	Rational Agents . . . . .	150
6.2.2	Agents Following the <b>GenIA<sup>3</sup></b> 's Default Design . . . . .	152
6.2.3	Agents Following the <b>GenIA<sup>3</sup></b> 's Extended Design . . . . .	155
6.3	General Results . . . . .	157
6.3.1	Agent's Mood Experience . . . . .	159
6.3.2	Agent's Decisions . . . . .	162
	<b>Conclusions</b>	<b>166</b>
<b>V</b>	<b>Conclusions</b>	<b>171</b>
<b>7</b>	<b>Conclusions and Future Work</b>	<b>173</b>
7.1	Conclusions . . . . .	173
7.1.1	Contributions . . . . .	175
7.2	Future Work . . . . .	178
7.3	Related Publications . . . . .	180
	<b>Appendix A: Ethical discussion</b>	<b>183</b>
A.1	Ethical discussion . . . . .	183
	<b>Appendix B: Affective Agents' Grammar and Examples</b>	<b>185</b>
B.2	Syntactic Specification of the Agent Language . . . . .	185
B.3	Syntactic Specification of the Multi-Agent System Language . . . . .	188
B.4	Source Code of Agents for Playing the Prisoner's Dilemma . . . . .	191
B.5	Source Code of Agents for Playing the Trust Game . . . . .	195
	<b>References</b>	<b>199</b>

Subject Index

221



---

---

## List of Figures

1.1	Snapshot of the <i>mWater</i> 's human-agents GUI for a Japanese auction. . . . .	8
2.1	Conceptual systems for a BDI affective agent. . . . .	36
2.2	Theoretical framework for intelligent affective agents. . . . .	38
2.3	Processes and cognitions of the LIDA architecture. . . . .	39
2.4	Processes and cognitions of the WASABI architecture. . . . .	42
2.5	Processes and cognitions of the BayesAct model. . . . .	44
2.6	Processes and cognitions of the emotion model of the infant robot. . . . .	46
2.7	Processes and cognitions of the formalization of OCC appraisal. . . . .	48
2.8	Processes and cognitions of the model of coping strategies. . . . .	49
2.9	Processes and cognitions of the Affective Belief Revision system. . . . .	50
2.10	Processes and cognitions of the computational model of empathy. . . . .	52
3.1	<i>GenIA</i> <sup>3</sup> . Sequences (solid line arrows), subprocess (dashed line arrows), and information exchange (dotted line arrows). . . . .	68
4.1	Extension of the reasoning cycle of AgentSpeak. . . . .	82
4.2	Simplified extension of the EBNF for the Jason Agent Language. . . . .	84
4.3	Simplified extension of the EBNF of the MAS project configuration. . . . .	87
4.4	Configuration of a Jason agent. . . . .	90
4.5	Code of the MAS configuration file for an agent 'player' following the default design. . . . .	107

## LIST OF FIGURES

---

5.1	Plays of the <i>betraying</i> agent against the <i>TESTER</i> agent in 100 rounds of the prisoner's dilemma. . . . .	117
5.2	Comparison of the years in prison of the <i>betraying</i> agent, the <i>FTFT</i> agent, and the <i>TESTER</i> agent playing each other 100 rounds of the prisoner's dilemma. . . . .	119
5.3	Extract of the output of 100 rounds of the PD played between the <i>playerAff</i> and the <i>TESTER</i> agents. . . . .	120
5.4	Simulation of the agent <i>trustor</i> in six rounds of the iterated TG with random back transfers. . . . .	125
5.5	Simulation of the agent <i>trustor</i> in six rounds of the iterated TG with descending back transfers. (a) and (c) trustor's NRO=[2,3]. (b) and (d) trustor's NRO=[6,7]. . . . .	126
6.1	Snapshot of the 3 <sup>rd</sup> game of the Black Jack experiment. The player has bet and he must decide between doubling, hitting, or standing. . . . .	131
6.2	Timeline of the Black Jack experiment by kind of treatment. Games can be preceded by an "aspiration" (triangles), an "autoefficacy" (stars), and/or a mood test (diamonds). . . . .	133
6.3	Classification of games and its assignation to specific kinds of game. H: hit, D: double, S: stand. . . . .	135
6.4	Difference of means between after and before a game grouped by <i>PA</i> and <i>NA</i> ("pos": positive group, "neg": negative group, "cont.win": control group when winning, and "cont.los": control group when losing). . . . .	140
6.5	Decision tree for classifying each individual's mood in positive (PosMood), negative (NegMood), or neutral (NeutMood). . . . .	144
6.6	Mean difference and confidence intervals with a confidence level of 95% in a pairwise comparison of bet by kind of mood. . . . .	145
6.7	Code of the plan for standing of the rational agents. . . . .	151
6.8	MAS project configuration for agents following the <i>GenIA</i> <sup>3</sup> 's default design. . . . .	153

---

## LIST OF FIGURES

6.9	An extract of the code of the agent following the <i>GenIA</i> <sup>3</sup> 's default design. . . . .	154
6.10	MAS project configuration for agents following the <i>GenIA</i> <sup>3</sup> 's extended design. . . . .	156
6.11	An extract of the code of the agent following the <i>GenIA</i> <sup>3</sup> 's extended design. . . . .	157
6.12	Comparison of the number of agents with positive (a), negative (b), and neutral mood (c) in games 6, 7, 16, 17 and 24 in the positive treatment. . . . .	158
6.13	Comparison of the number of agents with positive (a), negative (b), and neutral mood (c) in games 6, 7, 16, 17 and 24 in the negative treatment. . . . .	160
6.14	Means of <i>PA</i> for agents/participants in a positive (a), and negative treatment (b) in games 6, 7, 16, 17 and 24. . . . .	161
6.15	Means of <i>NA</i> for agents/participants in a positive (a), and negative treatment (b) in games 6, 7, 16, 17 and 24. . . . .	162
6.16	Percent of plays that match with the play suggested by the basic strategy in the non-training games. . . . .	163
6.17	Percent of the first plays 'double', 'hit', and 'stand' for each kind of simulation in game 7 for the whole sample. . . . .	164
6.18	Percent of the first plays 'double', 'hit', and 'stand' for each kind of simulation in game 7 by kind of mood. . . . .	165
B.4.1	Code of the MAS configuration file for the iterated prisoner's dilemma with two players: <i>TESTER</i> and <i>playerAff</i> . . . . .	192
B.4.2	Code of the agent <i>playerAff</i> in the iterated prisoner's dilemma. . . . .	193
B.4.3	Code of the agent <i>TESTER</i> in the iterated prisoner's dilemma. . . . .	194
B.4.4	Code of the agent <i>FTFT</i> in the iterated prisoner's dilemma. . . . .	194
B.4.5	Code of the agent <i>betraying</i> in the iterated prisoner's dilemma. . . . .	195
B.5.6	Code of the agent <i>trustor</i> in the iterated trust game. . . . .	196
B.5.7	Code of the agent <i>betraying</i> in the iterated trust game. . . . .	197





---

---

## List of Tables

2.1	The main processes and cognitions of emotion-cognition and emotion-behavior theories . . . . .	32
2.2	Summary of characteristics of the computational approaches. . . . .	55
3.1	Inputs and outputs for the processes of <i>GenIA</i> <sup>3</sup> . . . . .	72
4.1	Explanation of the main components of the EBNF extension for a Jason agent. . . . .	89
6.1	Results of T-test comparisons of the means after and before winning (or losing) by kind of treatment. . . . .	139
6.2	Comparison of the means of 11 samples of the positive, negative, and control groups with a confidence level of 95%. . . . .	141
6.3	Variables with better correlation and best model fitted for <i>vPA</i> and <i>vNA</i> . . . . .	142
6.4	Variables with better correlation and best models fitted for ‘bet of the game’. . . . .	147
6.5	Best predicting models of whether to hit, to stand or double by kind of mood and by kind of game. . . . .	148
6.6	Best predicting models of whether to hit or to stand by kind of mood and by kind of game. . . . .	149
6.7	Alternative model for predicting ‘bet of the game’ in the first non-training game without personality variables. . . . .	152



# Part I

## Introduction



---

## Introduction

The simulation of believable agents and virtual characters has been a challenge of artificial intelligence (AI) for decades. Research on Multi-Agent Systems (MAS) has traditionally focused on the search for rational solutions that maximize the quality or utility of the consequences of decisions, obtaining optimal solutions for problems of all kinds of complexity. To this end, computational agents have been provided with the capacity to reason, behave proactively, achieve goals, sense, and plan (among others). The most widely accepted model of rational agents is the Beliefs-Desires-Intentions (BDI) model, where beliefs, desires, and intentions are the core components of the agents' mental state. The BDI model (and architecture) is grounded on the philosophical theory of 'practical reasoning'. According to this theory, an agent decides, at each moment what actions it must perform in order to achieve its goals [Bratman, 1987].

Theory and technology on MAS have evolved in order to provide effective tools for simulating human-like behaviors of multiple agents, that are able to interact among them and with the environment. However, for human-like agents with affective characteristics and/or personality, this kind of approaches may not be the most appropriate. When we use the term *affective characteristics* we refer to a concept that includes the individuals' emotions, mood, feelings, personality,

# 1. INTRODUCTION

---

and related processes. Affective characteristics have been studied by experts on several disciplines, including psychologists, biologists, and neuroscientists, who often use either different terms to describe the same phenomena or the same terms to describe different phenomena. Our view of *emotions* is in line with the definition in [Ortony *et al.*, 1988]: “emotions are valenced reactions to events, agents, or objects, with their particular nature being determined by the way in which the eliciting situation is construed”. Moreover, we refer to *mood*, as the kind of affect of low intensity that is long-lasting with respect to emotions, that is not related to an event, agent, or object, and that changes according to emotions [Gohm *et al.*, 2003; Morris, 2012]. Our view of mood is in line with the definition of *core affect* in [Russell, 2003]: “a neurophysiological state that is consciously accessible as a simple, nonreflective feeling (...)”; In this work it is also used the term *affective state* or just *affect* to refer to an individual’s mood. On the other hand, *personality* is defined as “the dynamic and organized set of characteristics possessed by a person that uniquely influences his or her cognitions, motivations, and behaviors in various situations” [Ryckman, 2007], which is a definition that properly describes our view of personality in the present work.

Several disciplines, such as psychology, support the idea that affective characteristics have a strong influence on cognitive processes (e.g., the subjective evaluation of situations, or memory), and on decision-making. Psychological studies demonstrate that, when facing alternative choices, emotions guide decision-making towards an advantageous direction, influencing the subjective utility of the choices [Busemeyer *et al.*, 2007]. For example, an agent that imitates humans when playing a card game that involves gambling, needs mechanisms for making decisions about what move to make next, possibly biased by the current affective state of the agent, and also mechanisms for emotionally react to what happens in the game<sup>1</sup>. Studies also show that, in human-machine applications, the human-machine interaction is improved when virtual agents express emotions, enhancing human satisfaction [Klein *et al.*, 2002], and believability [Brave *et al.*, 2005; Picard & Liu, 2007], among others.

---

<sup>1</sup>Literature argues for the influence of emotions on decisions in this kind of games [Bechara *et al.*, 2000; Camerer *et al.*, 2011].

---

For example, an embodied virtual character that reacts emotionally to some external stimulus requires, not only an emotion-reaction mechanism, but also a mechanism for “interpreting” the stimulus and generating emotions. In neuroscience, methods such as neuroimages and unicellular activity records, have been used to develop a physical description of processes that are associated to emotions. Authors like Joseph E. LeDoux [LeDoux, 1998] and António Damásio [Damásio, 1994] made important neurological contributions by investigating the idea that emotions come from ancient parts of the brain that have evolved to generate appropriate responses to certain stimuli.

On the other hand, in experimental economics there is also enough evidence about the influence of subjective factors, such as emotions, on the decision-making process [Elbanna, 2006; Lerner *et al.*, 2004; Seo *et al.*, 2010]. Phenomena like the individual “risk aversion” (how tolerant is an individual to risk) [Loewenstein *et al.*, 2001], or the “illusion of control” (overestimation of individuals of their ability to control events) [Schwenk, 1984], also involve the affective characteristics of the decision-maker. Thus, for example, an owner of a property that has benefited from it for many years, and has affective bonds with it, would try to sell the property at a higher price than an investor that has not relation with the property. This is known in psychology and behavioral economics as “endowment effect” [Morewedge & Giblin, 2015]. By bearing all this in mind, one may assume that, when performing experiments with human that must make strategic decisions, it is very likely to get individual biased action profiles, regarding to what should be considered a rational solution. Therefore, when simulating human behavior in scenarios of complex decision-making, the representation of the individual affective characteristics may become an important parameter in order to get results better aligned with a real situation. These examples highlight the importance of affective characteristics for social and cognitive functions, and for decision-making, becoming indispensable characteristics for believable intelligent agents.

Computational intelligent affective agents are a suitable technology for creating simulations that cope with situations like those previously described. As it has been addressed by recent approaches (e.g., [Alfonso *et al.*, 2016a; Kefalas *et al.*, 2014; Pena *et al.*, 2011]), several applications can benefit from

# 1. INTRODUCTION

---

agents with affective characteristics, for example education, entertainment, disaster situations, training, and therapies, which have been improved through simulations with virtual humans. Also, the modeling of agents with an affective component as an element that influences their behavior, may improve the simulation of social interactions such as negotiation, and, in general, the process of decision making in complex situations [Bosse *et al.*, 2014].

## 1.1 Motivation

At the beginning of this research, believable agents were conceived by providing them with sophisticated mechanisms of reasoning and interaction [Alfonso *et al.*, 2011]. This new kind of believable agents was tested in the *mWater* case study [Alfonso *et al.*, 2014a; Giret *et al.*, 2011]. *mWater* is a virtual water-right market where trading negotiation is essential. It is embedded in a decision support system where water usage is subject to conflicts that may involve different types of negotiation. Trading negotiations that are performed in *mWater* follow different kinds of interaction protocols such as the face-to-face negotiation, the double blind auction, the Japanese auction, etc. As part of these negotiations several decisions should be taken. For example, at each step of a Japanese auction, participants should decide, in a limited period of time, whether to agree or not with a current bet (which increases at each round of bets).

The infrastructure of this market is a MAS where various technologies are involved. First, the MAS platform in itself, which manages agents and their interactions, allowing them to exchange information among them and also with the environment. Second, a language to define the agents behavior—in this case the Jason agent language [Bordini *et al.*, 2007]. Third, in order to support the interaction of humans with the software agents, a Graphical User Interface (GUI) was designed, as well as an artifact to orchestrate the communication between the GUI and the MAS. We used Magentix2 [Alberola *et al.*, 2008] as our MAS platform. Some of the reasons of selecting it included: i) it provides powerful techniques to facilitate agents' communication; ii) it supports interaction protocols between agents organizations/societies through



conversations' management; iii) it allows the use of high-level reasoning structures when programming the agents. Thus it provided proper technical tools for performing automated negotiations.

One of the interaction mechanisms provided by Magentix2 is the Conversation Factory [Fogués *et al.*, 2010]. Conversation Factories support FIPA interaction protocols, and each Conversation Factory allows to keep a complete interaction among two or more agents. On the other hand, Magentix2 supports high-level agent programming through Jason. Jason provides a kind of action called *internal action*. It is a structure that allows the execution of legacy code. By creating new internal actions, we allowed the Jason language became able to have access to the Conversation Factories provided by Magentix2 [Alfonso *et al.*, 2011] at a higher level of abstraction. These internal actions became part of the predefined internal actions of the Jason integration in Magentix2. By using some of these predefined internal actions, each agent could customize what it did in those steps of the conversations on which it needed to perform some 'reasoning' o take decisions, delegating details such as synchronization, timeouts, errors management, etc. to the platform.

One of the main goals of creating an automated trading market as *mWater*, was to perform simulations of real situations with agents representing human behavior. Because if this, one the greatest weaknesses of this approach was that the agent's 'reasoning' wasn't believable enough, and their decisions were never realistic enough. In order to have more realistic simulations, we created the above mentioned GUI, which allowed humans to interact with the software agents (see Figure 1.1). Nevertheless, without the human participation, results were far away from auctions in the real world. In these kind of interactions, such as auctions or, in general, negotiation, human's behavior is influenced by affective characteristics and emotional experiences [Picard, 2002; Zelenski, 2007]. For example when a participant performs a very high bet, this can produce anxiety, or losing the desired good can produce sadness. These affective states, in turn, can influence subsequent decisions. An approach that offered tools for a representation of more human-like behaviors, and more realistic results, was required. It was required an extensible, flexible, and open enough solution able to endow agents with affective characteristics. This solution

# 1. INTRODUCTION

---

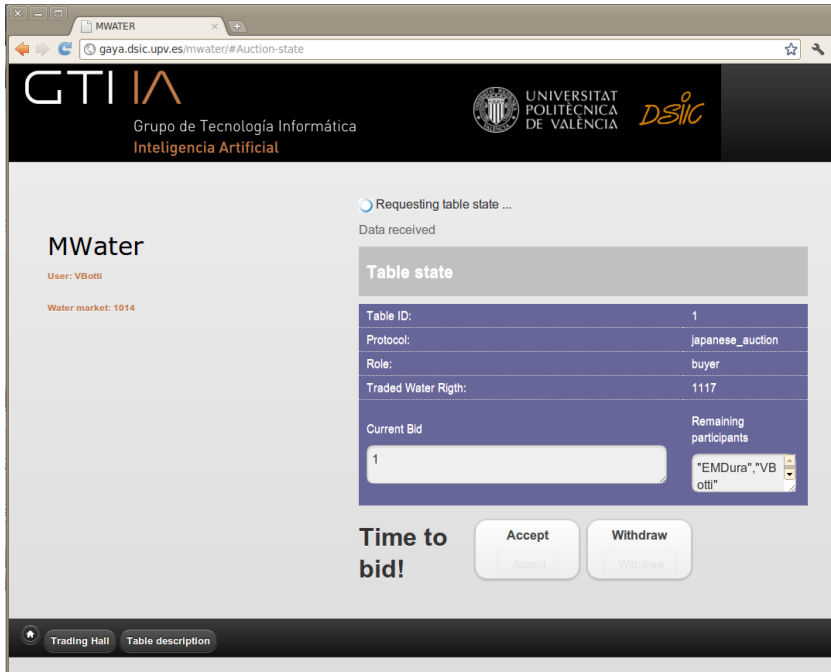


Figure 1.1: Snapshot of the *mWater*'s human-agents GUI for a Japanese auction.

should also be able to cope with the domain's requirements, and able to use different base theories in the search for the most credible behaviors.

## 1.1.1 Description of the Problematic Situation

Fields such as psychology, neuroscience, and economics have tried to understand the functioning of emotions and personality on individuals, as well as their influence in cognition and/or behavior. In the last decades, researchers of 'affective computing', which is considered by Rosalind W. Picard to be the "computing that relates to, arises from, or influences emotions" [Picard, 1997], have offered comprehensive solutions trying to address as many problems as possible when modeling affective processes, and they have also provided targeted solutions for specific application domains. These solutions are diverse in their definition, but share common basic representations and supporting

theories. Many computational approaches that model affective agents (mainly based on the cognitive perspective of emotions), model affect-related interconnected phenomena. The research literature proposes a wide amount of affective models that deal with some of: relationship between emotions and cognition, emotions and behavior, emotions and their evolutionary account, emotions for appraising situations, or emotion regulation (e.g., [Becker-Asano & Wachsmuth, 2010], [Hoey *et al.*, 2016], or [Franklin *et al.*, 2014]). These models are useful tools for addressing particular emotion-related issues. Furthermore, computational approaches that are based on particular psychological theories have also been proposed (e.g., [Steunebrink *et al.*, 2012]). They often address domain specific issues starting from a specific psychological theory. Nevertheless, when modeling a single affect-related phenomena, researchers have often to deal with either modeling all related processes (and hence making greater “unnecessary” efforts), or focusing on modeling the required phenomena, paying less attention to the rest of processes (and hence maybe missing important details). Therefore, there is a growing demand in computer science, of comprehensive approaches that promote standardization, and that favor an incremental development, the coordination in computer science (intradisciplinary coordination), as well as the exchange between computer science and other disciplines such as psychology (interdisciplinary exchange). This is a statement that, in the last decades, has been endorsed by relevant experts in the area (see [Reisenzein *et al.*, 2013] and [Hudlicka, 2014] for a reference). For example, Reisenzein *et al.* [2013] state that “by increasing the degree of modularization of architectures, it will be possible to build models out of components and thus avoid the constraints imposed by fixed architectures. Analogously, one may envision the development of a theoretical toolbox of basic theory-elements, formulated in a common language, from which theories of emotional agents (or of emotion modules for agents) can be constructed”.

Some examples of approaches (specifically agent architectures) that have been proposed for addressing these issues are the Methodology and Architecture for Modeling state and traIt effects on Decision-Making (MAMID) [Hudlicka, 2004, 2008] and Fearnot AffecTIve Mind Architecture (FAtiMA) Modular [Aylett & Paiva, 2012; Dias *et al.*, 2014]. Nevertheless, although MAMID is

# 1. INTRODUCTION

---

flexible regarding the level of parametrization, it is not so regarding the way these parameters influence cognition and reasoning. Besides, affect generation is limited regarding the process of emotion generation. FAtiMA Modular is more complete, being in line with the ‘modularized’ approach referred in [Reisenzein *et al.*, 2013].

## 1.2 Hypothesis

The guiding thread of this thesis work is defined by the following hypothesis:

A general purpose architecture for affective agents, with a default design, will allow the simulation of human affective traits and behavior in complex environments, performing better human-like simulations than the simulations of completely rational agents, and will also allow the incremental development of its individual components.

## 1.3 Goals

### 1.3.1 General Goal

In order to verify the hypothesis, this work aims to make an approach in line with the ‘toolbox of basic theory-elements’ referred in [Reisenzein *et al.*, 2013]. In this thesis, this toolbox is not only theoretical, but also practical, in that it includes (1) a general purpose affective agent architecture, with a default design, on the base of a general theoretical framework, (2) a language for implementing agents compliant to the proposed architecture, and (3) a platform implementing the default design of the architecture. This toolbox will allow building agents with both a rational and an affective component, able to have personality, emotions, and an affective state that may influence the agent’s decisions. Affective and rational processes, as well as affective and rational decisions, will be integrated in the platform in such a way that, alternative supporting theories can be adapted, and different relationships can be established between them. This approach fosters the common understanding

in computer science, the feedback between computer science, psychology and related disciplines, as well as the incremental research of these disciplines.

### 1.3.2 Specific Goals

In order to achieve the general goal, this work proposes the following specific goals:

- G1 Propose a **theoretical framework** for computational applications, which is the result of a horizontal review of the state of the art, mainly in psychology, that brings together the main concepts and processes regarding individual's affective characteristics and personality, as well as their influence on cognition and behavior.
- G2 Design a **general purpose affective agent architecture**, on the basis of the theoretical framework, that extends a traditional architecture of rational agents, and which includes a default design. This includes:
  - G2.1 Design a general purpose affective agent architecture where different psychological theories (relative to individual differences, affect generation, affect dynamics, and affect influence on cognition and behavior) can be implemented, on the basis of a common conceptual system.
  - G2.2 Integrate affect-related processes in the BDI architecture of rational agents in a semi formal way.
  - G2.3 Extend a formal language of rational agents (Jason), for it to be compliant with the proposed architecture.
  - G2.4 Propose a default design inspired by widely accepted theories for affect generation, affect experience, and for the influence of the affective state on behavior.
  - G2.5 Facilitating to set an equilibrium between the rational and the affective sides of the agent, by offering means to adjust the level of rationality of an individual, and the frequency of rational and affective processes.

## 1. INTRODUCTION

---

- G3 Extend the Jason's **MAS platform** for creating agents with the proposed architecture.
- G4 Perform an **experiment with real participants** in order to gather information regarding how participants' affective state changes after an event with emotive content, and how participants make decisions in relation to their affective state.
- G5 **Verify whether the hypothesis of this work is fulfilled by comparing**, under the same experimental conditions, the results of the execution of agents following the proposed architecture with (1) results of classical problems, and with (2) results of the experiment with real participants in three directions:
- G5.1 Verify that real situations can be represented with the architecture's default design.
  - G5.2 Verify that this default design can be extended in order to fulfill specific requirements.
  - G5.3 Verify that the results obtained through the affective MAS platform are more realistic than the results obtained by rational approaches.

### 1.4 Structure of the Document

In this chapter, an introduction to the thesis work has been offered, which has outlined its motivation, its hypothesis, and its general and specific goals. The rest of the work is organized as follows.

**Chapter 2** explores the main processes and concepts, mainly of psychology, but also of neurology, by performing a horizontal review of the state of the art, regarding individuals' affective characteristics and personality. A theoretical framework that integrates those processes and concepts is also proposed. Besides, examples of how both, specific purpose, or general purpose computational approaches have included those processes and concepts, are also described.

**Chapter 3** proposes the *GenIA*<sup>3</sup> architecture, by describing its main components and structure, as well as by describing how practical reasoning and affective processes and characteristics are integrated on it. A description of the main attributes that an agent should possess according to *GenIA*<sup>3</sup>'s structure is also offered.

**Chapter 4** proposes an extension of the Jason agents' reasoning cycle in accordance with *GenIA*<sup>3</sup>, as well as the transition rules for the new steps through the extension of the AgentSpeak's operational semantics<sup>1</sup>. This chapter also proposes the default design of *GenIA*<sup>3</sup>.

**Chapter 5** proposes two case studies, where the classical games 'Prisoner's dilemma' and 'Trust game' are used for offering alternative solutions to rational solutions of experimental economics through *GenIA*<sup>3</sup>.

**Chapter 6** describes the experiment performed with real participants, which play several rounds of the BlackJack game. This chapter also describes three simulations: one of fully rational agents, another of agents compliant with the default design of *GenIA*<sup>3</sup>, and a simulation with agents compliant with an extended design of *GenIA*<sup>3</sup>. General results and some conclusions are offered at the end of the chapter.

**Chapter 7** summarizes the conclusions of this thesis work as well as its main contributions. Open questions and future related lines of research are also addressed in the section of future work. Finally, the chapter offers a relation of the bibliographical work that has been published during the development of this thesis.

---

<sup>1</sup>Jason implements the operational semantics of the AgentSpeak agent language [Vieira *et al.*, 2007].





## Part II

# State of the Art



---

## State of the Art

### 2.1 Introduction

When modeling intelligent and social agents, two main issues spring to mind: the agents' practical reasoning and the agents' interactions. This chapter proposes the affective processes that are related to both as well as how they are related. To this end, a horizontal review of core philosophical and psychological concepts and theories are reviewed. Considering that agents are situated in environments where they interact with similar but not identical entities, the chapter also proposes the personality factors that need to be considered in order to establish differences between agents, their behavior, and reasoning. Specifically, this chapter aims to achieve the Goal G1 of this thesis (see Chapter 1), by proposing a theoretical framework for computational applications, that brings together concepts and processes regarding individuals' affective characteristics and personality, as well as their influence on cognition and behavior.

Our view of the relationship between emotions and practical reasoning is in line with what contemporary philosophy proposes:

“Contemporary philosophy of emotion attempts something stronger, however, in according emotions a role in practical reasoning. Mak-

## 2. STATE OF THE ART

---

ing this an integral role—understanding emotions as functioning within practical reasoning rather than just as spurs to it—means interpreting emotions in normative terms, as providing or expressing potential reasons for action, and as themselves subject to rational assessment and control, contrary to the traditional view of emotions as ‘passive’ phenomena” [Greenspan, 2004].

In this view, emotions are very linked to rationality, which is a view shared by research areas such as psychology and neuroscience. For example, Antonio Damásio’s neurological studies [Damásio, 1994] showed that emotions played an important role in rational thought, unlike Descarte’s philosophy, which separates the mind from the body. He used a real story about a man called Phineas P. Gage who had suffered an accident where an iron rod passed through his head causing a wound in the ventromedial prefrontal cortex. After the accident, the man exhibited a change in his personality and behavior, exhibiting, (in appearance), a lack of emotions and feelings. With this analysis, Damásio demonstrated the link between the physical and feeling and how important emotions are in rational thought and decision-making.

In this chapter we first make a review of the psychological literature that is relevant from a computational perspective (some of which is supported by neurological findings) in order to describe how practical reasoning and affective features are related. The theories that allow an individual to be differentiated from others are presented later. Then, the affective features/processes that are related to agent cognition and practical reasoning are presented, as well as examples of computational approaches that model specific affective phenomena in agents with practical reasoning. Finally, a comparison of approaches that are similar to the one proposed in this thesis is performed.

### 2.2 The Influence of Emotions in Cognition and Practical Reasoning

Philosophical, psychological, and neurological studies argue for the relation between emotions and cognition or between emotions and practical reasoning.

## 2.2. Emotions in Cognition and Practical Reasoning

---

For example, emotions may influence memory enhancement and the evaluation of the situation. Emotions also play an important role in social interactions, or help to narrow the set of possible options in decision-making [Greenspan, 2004; Kirman *et al.*, 2009; Levine & Pizarro, 2004]. Emotions “serve as a powerful organizing force, not just for behavior, but for perception, judgment, and memory” [Levine & Pizarro, 2004]. Having been inspired by philosophical and psychological studies on affect-related issues and their relationship with rationality and cognition [Greenspan, 2004; Kirman *et al.*, 2009; Levine & Pizarro, 2004], we have made a selection of affective phenomena, each of which are described in Sections 2.2.1 to 2.2.7. These sections mainly focus on the relation between these affective phenomena and cognition and these affective phenomena and behavior, by emphasizing the changes produced in the individual informational state (represented by beliefs), motivational state (represented by desires), and deliberative state (represented by intentions). Some of the descriptions of these relations are also supported by neurological theories.

### 2.2.1 Emotions and Memory

There is evidence that people better remember events that have been emotionally significant (i.e., those with more vividness) than non-emotional events [Levine & Pizarro, 2004]. Studies have shown several results related to what information is remembered and how it is remembered as well as whether or not emotional memories are indelible. It has been traditionally believed that highly emotional events lead to indelible memories [James, 1890; LeDoux, 1992], although other results have demonstrated that these memories may not necessarily be accurate or consistent [Talarico & Rubin, 2003] or that they may even fade over time. In [Bower *et al.*, 1981], it is suggested that this happens because individuals’ information encoding and retrieval is performed by matching this information with the corresponding elicited emotional state. These facts confirm that the primary function of memory is to guide future behavior in order to either avoid or promote familiar situations, without the need of keeping an exact record of past events [Levine & Pizarro, 2004].

## 2. STATE OF THE ART

---

### 2.2.2 The Influence of Emotions on Behavior

An important and influential contribution regarding the influence of emotions on behavior was the one made by Damásio [Damásio, 1994]. He stated that there is a strong link between the physical and feeling and that emotions are significantly important in rational thought and decision-making. Damásio proposed the definition of the *somatic marker*, which is a mechanism whereby emotions guide behavior and decision-making. This definition is contrary to Descarte's view of pure reason, which considers emotions to be more of a hindrance than an aid for reasoning.

The way emotions and affective states determine or influence behavior has been widely discussed. One of the debates has been centered on whether emotions exert a direct causal influence on behavior, or, on the contrary, if they are an input into the processes of decision and behavior regulation as a feedback system [Baumeister *et al.*, 2007]. In line with these ideas, Damásio proposed the concepts of *primary* and *secondary emotions*. He stated that *primary* emotions were innate or “wired in at birth” and dependent on “limbic system circuitry, the amygdala and anterior cingulate being the prime players”; they allow having reactive responses in the case of immediate danger (which is a direct causal influence on behavior). If primary emotions become conscious, then this can make the response more flexible also allowing past interactions with the environment to be considered. On the other hand, *secondary* emotions (also known as social emotions), are conscious and more elaborated (e.g., jealousy, guilt, or pride) [Damásio, 1994]. According to [Baumeister *et al.*, 2007], “conscious emotion commands attention and stimulates analysis, learning, and adaptation”. One way of describing this conscious influence of emotions is through the principle of learning by reinforcement, where positive or negative affects are associated to similar events depending on past outcomes for these events [Gollwitzer, 1999]; hence, future behavior is guided by anticipated emotional outcomes. Also, experiments in [Demaree *et al.*, 2009] demonstrate that a specific dimension of the affective state of an individual is a significant indicator of the level of risk that is associated to the decisions that are made.

Another theory that describes how emotions guide behavior, specifically in

## 2.2. Emotions in Cognition and Practical Reasoning

---

social interactions, is the social psychological *affect control theory*. It proposes that human decisions (as well as perceptions and emotional experiences) respond to the need of minimizing deflections between sentiments about social situations that are culturally shared and the evaluations resulting from interactions in those situations [Heise, 2007]. For example, when a professor decides to advise a student instead of yelling him, he is behaving according to the cultural expectations of professorial behavior. This allows building models and simulating interactions that can be both goal-directed and affect-sensitive [Hoey *et al.*, 2016].

### 2.2.3 Emotions and their Evolutionary Account

Two important exponents of the evolutionary account of emotions are Charles Darwin and William James. In his work titled “The Expression of Emotion in Man and Animals” [Darwin, 1872], Darwin states that emotions are a survival function (that comes from an evolutionary process) to solve the problems we have as species. This is why there are similarities between our emotions and those of closely-related species. This work focuses on emotional *expression* and shows a relation between the movements and facial expressions of humans and those of other animals. In his theory, Darwin insists on the idea that emotions have an important function and therefore a survival value. On the other hand, James defined emotions as the feeling that arises from bodily changes following the perception of an exciting fact [James, 1884]. He stated that bodily changes come first, and, after them, we experience emotions in such a way that it would be impossible to have emotions without bodily changes. As Darwin did, James considered emotions to be a survival factor. The Jamesian theory had many followers, but many people also supported the opposite position. For example, authors like Walter Cannon [Cannon, 1927], thought that we feel emotions and experience physiological reactions such as sweating or muscle tension simultaneously in such a way that emotions arise when the thalamus sends a message to the brain after perceiving a stimulus. This results in a physiological reaction which is known as the “Cannon-Bard” theory.

Other researchers have followed these evolutionary perspectives by dis-

## 2. STATE OF THE ART

---

agreeing to a greater or lesser extent with Darwin's point of view. This is the case of William McDougall, Robert Plutchik, Paul Ekman, Carroll Izard, and Sylvan Tompkins (among others). Some of those authors [Ekman *et al.*, 1987; Fridlund, 1994; Russell, 1994] have identified a set of fundamental or primary emotions that are common in every species. These emotions represent survival patterns of responses to events in the history of evolution, and they play an "adaptive role in helping organisms deal with key survival issues posed by the environment" [Averill, 1980]. Based on this, a set of universally recognized facial expressions has been identified, which may vary in number, but that maintains the idea that emotions respond to an evolved adaptive response pattern. A set of "basic-level emotion categories" that surpasses cultural barriers has also been identified [Shaver *et al.*, 1987, 1992].

The evolutionary account of emotions also has neurological roots [Damásio, 1994; LeDoux, 1998]. For example, in [LeDoux, 1998], LeDoux tries to explain that emotions are part of a complex neural system that has evolved in order to survive. Therefore, the emotional responses are hard-wired into the brain's circuitry, which is built through learned experience. He explored the mechanisms that produce emotions underlying the brain, and he found that, for example, the brain can detect danger even before experiencing fear, leading to physiological responses like sweaty palms or muscle tension.

### 2.2.4 Emotions in the Evaluation of the Situation

It is considered that emotions consist of multiple components [Moors & Scherer, 2013]. These components include the following: the *cognitive* component, which involves evaluations of the person-environment interaction; the *motivational* component, which includes forms of action readiness; the *somatic* component, which includes physiological responses that are peripheral; the *motor* component, which deals with instrumental and expressive behavior; and the *subjective* component, which involves subjective experiences. The *cognitive* component is underpinned in the cognitive theory, which is the one that deals in great detail with how the "perception of the exciting fact" mentioned by James [James, 1884] becomes an emotion. This was called the "appraisal" process by authors



## 2.2. Emotions in Cognition and Practical Reasoning

---

like Magda Arnold [Arnold, 1960]. There are coincidences as well as differences in the content and number of appraisals (processed appraisal variables) that are made. Theories generally agree on appraisal variables like goal relevance, goal congruence, certainty, coping potential, and agency (whether an event is caused by oneself, other people, or external circumstances) [Moors *et al.*, 2013]. Other appraisal variables are not considered in all theories (e.g., novelty, pleasantness, type of goal, or norm compatibility) [Moors *et al.*, 2013]. Theoretical and experimental research on appraisal has deepened the study of the relation between patterns of appraisal values and emotions, action tendencies, physiological responses, and facial or vocal expressions. Some key questions include whether a representation of emotion must be performed, how it must be performed, and how appraisal patterns influence other components [Moors *et al.*, 2013].

For example, in [Scherer, 1987], Scherer proposed a component process model of appraisal where appraisal consists of an invariant sequence of “stimulus evaluation checks”. Specifically, an emotion is defined as “an episode of interrelated, synchronized changes in the states of all or most of the *five organismic subsystems*<sup>1</sup> in response to the evaluation of an external or internal stimulus event as relevant to major concerns of the organism” [Scherer, 1987, 2001]. The possible states of these subsystems are the components of an emotion episode [Scherer, 2005]. The role of *concerns* stands in all appraisal theories. They can be defined as individual needs, ideals, values, norms, and all the individual cares about [Frankfurt, 1988; Frijda, 2007; Lazarus, 1994; Scherer, 2004].

Another important appraisal theory is the one proposed by Ortony, Clore, and Collins in their work titled “The Cognitive Structure of Emotions” [Ortony *et al.*, 1988]. One of the most important contributions of their work is that the authors specifically developed their theoretical approach with the aim of implementing it on a computer. This has been the most widely accepted and used proposal by researchers seeking to work in this direction. The authors propose a model of emotions that they called “OCC” (the first letters of the

---

<sup>1</sup>The five organismic subsystems are: Information processing, Support, Executive, Action, and Monitor. See [Scherer, 1987] or [Scherer, 2001] for a more detailed description.

## 2. STATE OF THE ART

---

author's names). This model not only proposes a classification of emotion types by considering aspects of the situation appraised, but it also proposes quantitative aspects of emotions. Variables such as *potentials*, *thresholds*, and *intensities* are used to describe these quantitative aspects.

Richard Lazarus' work [Lazarus, 1994; Smith & Lazarus, 1990] differentiates "primary" and "secondary" appraisal, which may or may not be sequential. He stated that, in a "primary" appraisal, an event is evaluated to determine whether it helps or hinders the achievement of a goal, which fosters or inhibits the desires of an individual. In a "secondary" appraisal, an individual assesses its capabilities and resources to deal with an event. On the other hand, in Roseman's work [Roseman, 1996; Scherer *et al.*, 2001], appraisal patterns are considered to be dimensions. The author distinguishes seven appraisal dimensions, including some concepts such as the unexpectedness of an event, its probability, control potential (referring to an individual's capacity to cope with an event), and agency (i.e., responsibility regarding the event).

Some theories, such as the "social constructivist", argue that emotions come from cultural facts, and their meaning and coherence are given by learned social rules. According to Averill [Averill, 1980], emotions can only be analyzed on a social level. They play an important social role at both the interpersonal and social level. Thus, for example, the fact of being wronged may vary among different cultures because it means that a social rule has been violated, which can be different depending on what it means to each individual. According to the social constructivist perspective, if the appraisal that generates emotions is a biological adaptation, then its content is cultural.

### 2.2.5 Emotion Regulation

Emotion regulation is the process whereby actions are performed to reach a desirable affective state, in either a conscious or unconscious way, by using indirect individual coping strategies that are oriented to change their affective state [Gross & Thompson, 2011; Larsen, 2000; Thayer *et al.*, 1994]. This doesn't imply that people can directly control moods and emotions, but they can attempt to do so. These strategies may involve maintaining or altering

## 2.2. Emotions in Cognition and Practical Reasoning

---

beliefs, desires, intentions, or expectations that motivated the corresponding actions through the appraisal process. For example, these strategies can include suppressing information when confronting with an event with low controllability and/or negative impact or focusing attention away from the emotional event. Strategies can also include increasing the probability of a pending outcome that is desirable through wishful thinking [Marsella & Gratch, 2009; Smith & Lazarus, 1990]. According to [Lazarus & Folkman, 1984], two kinds of coping (as ways of dealing with stressful situations) can be distinguished: *problem-focused* and *emotion-focused*. *Problem-focused* coping generally takes place when there is a possible solution to the problem that originated the change in the affective state, and a behavior oriented toward the management of the problem is created. *Emotion-focused* coping is more probable when the individual believes that there isn't a solution to the problem, and the behavior is oriented toward the regulation of the corresponding emotional response.

Psychologists like Sigmund Freud proposed a similar theory called “defense mechanisms”. These are individual resources for preventing the *ego*<sup>1</sup> from suffering danger (e.g., anxiety) [Freud, 1937]. They are activated unconsciously and some of their categories are: *repression* (the *ego* attempts to keep the undesirable impulses of the *id* in the unconscious); *denial* (when the person refuses to perceive unpleasant information from the environment); *sublimation* (when the unacceptable impulse is transformed into a socially acceptable, even productive form); *projection* (the tendency to attribute our own unacceptable desires to other people), etc. [Ryckman, 2007].

Similarly Carl Rogers argued that when experiencing incongruence between the ideal and the actual self<sup>2</sup>, the individual becomes anxious [Rogers *et al.*,

---

<sup>1</sup>According to Freud, individual instinctual needs set the starting point for the psychic energy that is available. The elements that compete for it are the *id*, the *ego*, and the *superego*. They are systems of the mind that interact dynamically to influence behavior. Freud defined the *id* as the original aspect of personality, which is located in the unconscious part of the mind; it contains primitive desires such as unconscious sexual and aggressive instincts. The *ego* and the *superego* lie in the conscious, unconscious, and preconscious part of the mind. The *ego* provides realistic direction for the impulses and partially carries out the aims of the *id*; the *superego* describes the internalization of social values in the individual. It interacts dynamically with the *id* and the *ego*.

<sup>2</sup>The ideal self is what you want to be and the actual self is what you really are now.

## 2. STATE OF THE ART

---

1989]. To defend against this, Roger thought that we use defenses that can include the distortion of experience in order to make something acceptable or to prevent threatening experiences from reaching awareness at all.

### 2.2.6 Evaluative Implications of Emotions

Each individual may have different ways of interpreting the strength and content of what they sense, and this process may be influenced by the current affective state. According to [Frijda *et al.*, 2000], emotions influence the “content and the strength of an individual’s beliefs and their resistance to modification”. It is known that, for example, when we are in a negative affective state, we question our beliefs more than in a neutral affective state, and we are more prone to accept new information. Also a positive affective state makes us to rely more on our current beliefs [Pimentel & Cravo, 2005]. Some other works support these findings [Bodenhausen *et al.*, 1994; Forgas, 1998]. For example, in [Bodenhausen *et al.*, 1994], it is shown that when happy people perform judgments, they are more influenced by stereotypes (i.e., their own beliefs and mental shortcuts about things) than people in a neutral mood. Another example is the act of self-deception, which is commonly known as the internal biasing processes that takes place when we adopt a desired belief in the face of possibly contradictory evidence<sup>1</sup> [Demos, 1960; Ellsberg, 1961].

### 2.2.7 Emotions and Social Interaction

Emotions greatly help in the understanding of social interactions. When interacting with other individuals, factors like empathy, emotional contagion, or social learning may take place [Kirman *et al.*, 2009], and hence may generate emotional experiences. Some of these phenomena are grounded on the *theory of mind* (ToM)[Whiten, 1991], which studies the human ability to attribute mental states (like beliefs, desires, or intentions) to other people, and to use those mental states to reason or to feel like others would do. Besides, the capacity of individuals to identify emotional signals in facial expressions or body gestures allows these phenomena to take place (see e.g., [Hess *et al.*, 1998; Williams,

---

<sup>1</sup>Assuming that contradictory evidence leads to negative affective states.

2002]). *Affective empathy* can be described as the process whereby a perceiver experiences an affective state that matches the affective state perceived in another individual (for instance see [De Waal, 2008]). On the other hand, *emotional contagion* is regarded as “The tendency to automatically mimic and synchronize expressions, vocalizations, postures, and movements with those of another person and, consequently, to converge emotionally” [Hatfield *et al.*, 1992, p. 153]. Nevertheless, in works like [Hess & Fischer, 2014], *contagion* is demarcated from *mimicry* by referring to contagion as the “matching of a subjective emotional experience”, while mimicry refers to the “matching of nonverbal displays”.

## 2.3 Interindividual Differences

Individuals behave according to specific traits, depending on individual characteristics or other factors like their culture. Hence, the emotional experience, the relation between affect and cognition, and the relation between affect and reasoning vary from individual to individual. Considering this fact, the representation of these features in artificial intelligence, specifically in believable intelligent affective agents, becomes essential in order to highlight their singularities [Castelfranchi *et al.*, 1998]. According to [Rizzo *et al.*, 1997], “among the most important features of a believable agent are a marked personality and emotions”.

Personality is considered to represent the long-lasting individual characteristics that influence motivations and behaviors when facing a given circumstance [Howard & Howard, 1995; John & Srivastava, 1999; Santos *et al.*, 2009]. Researchers of the different perspectives of personality have focused on both describing the personality dynamics within an individual and determining its structure (i.e., the individual’s traits that differentiate him or her from the rest). Approaches that are focused on personality dynamics are concerned with the *intrapersonal* (also called *intraindividual*) structure and mechanisms of personality and how they evolve over time taking into account the individual’s interaction with the environment, his or her goals, motives, and beliefs. Structural approaches (also called *interindividual*) are more concerned with

## 2. STATE OF THE ART

---

individual differences determined by trait constructs [Read *et al.*, 2010]. The analysis of the *intraindividual* structure of personality is out of the scope of this thesis since determining how personality is shaped over long periods of time is not of interest for this work. This work focuses on the *interindividual* structure instead, which allows establishing individual characteristics for a computational simulation that is time bounded.

In theories that study the *interindividual* structure of personality, some researchers tried to find a relationship between physical characteristics and personality traits: for example, the relation between psychological characteristics and facial features such as beauty [Baron *et al.*, 2006], or the shape and contours of the skull [Davies, 1955] (also called phrenology). Even a classification for body builds was made by William Sheldon [Sheldon *et al.*, 1954]. All those theories were too subjective and imprecise. Other accepted theorists in this field were Gordon Allport, Raymond Cattell, and Hans Eysenck. They studied personality from a *traits* perspective. Gordon Allport, who is considered to be the father of personality theory, defined a trait as “a generalized and focalized neuropsychic system (peculiar to the individual) with the capacity to render many stimuli functionally equivalent and to initiate and guide consistent (equivalent) forms of adaptive and expressive behavior” [Allport, 1937]. He found more than 4000 words in an English-language dictionary that describe personality in a total of nearly 18000 words [Allport *et al.*, 1936]. He also categorized the personality traits into three levels [Allport, 1961]:

**Cardinal Traits:** Those dominant and pervasive characteristics in a person’s life. People with these characteristics are often known by their traits, even in their names.

**Central Traits:** These traits are those that control the person’s behavior to a lesser degree, so they do not possess the dominance of a cardinal trait. Nevertheless, they are important. These are major characteristics that we often use to describe another person.

**Secondary Traits:** They often appear under a specific circumstance and they are related to attitudes or preferences. They are peripheral to

a person.

The theorist Raymond Cattell reduced Allport's initial list of words that describe personality down to 171 words. This was done mostly by eliminating synonymous words and by combining common characteristics. Then, by using the statistical technique of factor analysis and other data collection techniques, he identified just 16 key personality traits [Cattell & Tregaskis, 1965]. These traits would be useful to explain personality functioning according to Cattell. He also created an *Econetic Model* [Cattell, 1979]. With this model, Cattell wanted to remedy the deficiency of trait theories since they did not consider the role of the environment in predicting behavior. Therefore, this model postulates the interaction between traits and physical, social, and cultural environments and their influence on behavior.

Hans Eysenck defined a typology that is hierarchically organized. It consists of three levels: *types*, *traits*, and *habits* [Eysenck, 1970]. The *types* level is the most abstract and is based on the intercorrelations among traits (like sociability, impulsivity, or liveliness). The intercorrelations among habitual responses (e.g., having breakfast immediately after waking up) is what he called *traits*. Finally, *habits* are inferred from observable specific responses. Eysenck believed that people could be described in terms of three supertraits with a genetic basis, which are considered to be the ones that best describe the functioning of the personality [Eysenck, 1982; McCrae & John, 1992]. These supertraits are: introversion-extraversion (a person that rates high in introversion directs his or her attention to inner experiences, while extroverts focus their attention on the environment or other people), stability-neuroticism (neuroticism refers to the tendency of the individual to become upset or emotional, and a stable individual is emotionally constant), control-psychoticism (psychotic individuals differ from neurotics in the severity of their conduct disorders, like insensitivity to others, hostility, or manipulation of others, and a controlled person does not have these characteristics).

Stimulated by the works of Eysenck and Cattell, other researchers tried to find middle ground between the two studies. They believed that Cattell focused on too many traits and Eysenck on too few [McCrae & Costa, 1997]. As a

## 2. STATE OF THE ART

---

result, they derived five relatively strong factors for describing the personality of each individual [Digman & Takemoto-Chock, 1981; Goldberg *et al.*, 1990; Norman, 1967; Tupes & Christal, 1992]. The exact label for each dimension often varies from one researcher to another. In [Goldberg *et al.*, 1990], Goldberg proposed a five-dimension classification for personality: *Extraversion*, *Agreeableness*, *Conscientiousness*, *Neuroticism*, and *Culture*. Nevertheless, the most common denomination is the one proposed in [McCrae & John, 1992]. The authors describe the traits as follows. *Openness* is related to culture, intellect, appreciation for art, adventure, curiosity, originality, or unusual ideas. *Conscientiousness* is related to the will to achieve something, discipline, planned behavior instead of spontaneous behavior, organization, or responsibility. *Extraversion* is related to energy, positive emotions, search for the company of others and stimulation, a talkative individual. *Agreeableness* is when an individual is cooperative and compassionate instead of suspicious or antagonistic, a generous, kind, and forgiving individual. *Neuroticism* is the tendency to experience negative emotions such as anger or depression. An individual with a high level of neuroticism is an anxious, self-pitying, and unstable individual.

Besides adepts, the theory of traits has many critics [Briggs, 1989; John, 1989; John & Srivastava, 1999]. This is mainly due to its weakness in that it cannot represent a taxonomy of traits that is really comprehensive. Another commonly mentioned problem is that traits are often poor predictors of behavior since an individual does not always behave as expected according to his or her predominant traits. Also, this theory does not address how individual differences in the development of personality emerge. Some researchers state that more than five factors are needed to describe personality appropriately [McCrae & John, 1992]. Nevertheless, it is “an important breakthrough in the study of personality” [Ryckman, 2007].

### 2.4 Affective Agents with Practical Reasoning

The affect-related cognitive theories described in Section 2.2 help to propose domain-independent, basic computational structures and processes that an in-



telligent affective agent should have. So do general architectures for autonomous agents. Specifically, this work uses the BDI (Beliefs-Desires-Intentions) agent architecture, because of its suitability to create computational models of emotions [Reisenzein *et al.*, 2013]. In this section, the concepts addressed by the theories of Section 2.2 are summarized. Then, the links between those concepts and the BDI's cognitive processes and components are established through a conceptual system. Finally, a set of computational approaches (that address, to a greater or lesser extent, each one of the relationships between emotions and cognition, and between emotion and behavior of Section 2.2) are compared, by using the proposed conceptual system.

### 2.4.1 Conceptual Systems for Intelligent Affective Agents

Table 2.1 summarizes the main processes described in Section 2.2, which are grouped by emotion-cognition or emotion-behavior relations. The table also shows the cognitions involved. In the column ‘Cognitions involved’ of Table 2.1, we use ‘affective state’ as a more generalized concept that includes one or more aspects of the agent affective state (such as emotions or mood). Table 2.1 shows that in the relation of emotion and memory, two important processes can be highlighted: 1) the process in charge of determining what and how *emotional memories* are stored as well as the *event* associated to the process; and 2) the process that determines how these memories evolve over time. On the other hand, according to research on the influence of emotion on behavior, two broad processes are depicted in Table 2.1. The first one manages the direct influence of emotion on behavior (the process of generation of reactive behavior), and the second uses emotions as input for decision-making (the process of decision-making). Both forms of behavior generate or modify the agent’s *intentions*. Researchers agree that the generation of reactive behavior is mainly done on the basis of *past experiences* and *emotional memories*. The same applies to decision-making, which also takes place by mainly evaluating *expectations* and *culturally shared evaluations*. On the other hand, theories on the evolutionary account of emotions, mainly address behavior that is related to physical manifestations such as bodily changes, body gestures, or the

## 2. STATE OF THE ART

---

Table 2.1: The main processes and cognitions of emotion-cognition and emotion-behavior theories

<b>Relations emotion-cognition /emotion-behavior</b>	<b>Main processes</b>	<b>Cognitions involved</b>
Emotions and memory	Storing of emotional memories Maintenance of emotional memories	emotional memories (emotional events)
Influence of emotions on behavior	Generation of reactive behavior Decision-making	intentions, past experiences, emotional memories, culturally shared evaluations, expectations, affective state
Emotions and their evolutionary account	bodily changes, generation of facial expressions, body gestures, etc.	—
Appraisal	Perception  Affect generation	percepts, beliefs, probability of beliefs (herein ‘belief probability’), concerns, desires, norms, capabilities, resources, social rules, affective state
Regulation of emotions (herein Emotion regulation)	Generation of coping behavior	beliefs, desires, intentions
Evaluative implications of emotion	Perception	percepts, beliefs content and strength
Emotions and social interactions	Affect generation (through empathy, emotional contagion, or social learning)	Beliefs, Affective state

## 2.4. Affective Agents with Practical Reasoning

---

generation of facial expressions. In general, these theories do not emphasize those changes in individuals' cognitions that have evolutionary explanations. The next emotion-related process in Table 2.1 is the one related to appraisal, which involves the processes of perception and affect generation. According to the different appraisal theories analyzed in Section 2.2.4, we have identified the cognitive components that are generally used in the appraisal process. Some examples are: *percepts*, *beliefs* (as well as the values of associated *probabilities*), *concerns*, *desires*, *capabilities*, *resources*, etc. The appraisal process is the main influence on the *affective state* of the individual. Emotion regulation, on the other hand, implies the generation of coping behavior. Corresponding theories argue for the influence of coping behavior on *beliefs*, *desires*, or *intentions*. The evaluative implications of emotions mainly involve the process of perception. This process evaluates *percepts* and determines the *content* and *strength* of beliefs. Finally, Table 2.1 shows another process of affect generation that is originated by the relations of emotions and social interactions described in Section 2.2.7. In this process, the *beliefs* about other agents and about the situation also play an important role.

### 2.4.1.1 Cognitive Processes of Intelligent Agents

According to [Weiss, 1999], general architectures for intelligent agents can be classified into four classes of agents: *logic-based* (decisions are taken through logical deduction), *reactive* (decisions respond to direct mappings of situations to actions), *belief-desire-intention* (beliefs, desires, and intentions are manipulated in order to take decisions), and *layered* (where different software layers, which reason about the environment at different levels of explicitness and abstraction, determine the decisions to take). Specifically BDI (Beliefs-Desires-Intentions) architectures are grounded on the philosophical roots of *practical reasoning*, which is the process whereby, at each moment, an agent decides the actions to perform to reach its goals [Bratman, 1987]. The BDI architecture is well-suited for building computational models of affective agents. Some of the arguments that support this include that it has its foundation in philosophical and psychological theories, it contains the main components of

## 2. STATE OF THE ART

---

emotions (which are beliefs and desires), and it can be enhanced (along with its logic), extended and/or combined to account for new structures or processes [Reisenzein *et al.*, 2013].

Practical reasoning involves mainly two processes: *deliberation* (where the agent decides what to do by determining the goals to be attained) and *means-end* reasoning (which involves how to achieve these goals). Most BDI agents perform these processes through four more specific processes which are: *belief revision* (which includes *perception*, and determines new *beliefs* on the base of percepts and current beliefs); *option generation* (which determines the agent *desires*, on the basis of its *beliefs* and current *intentions*); *filter* (which determines new *intentions* on the basis of current *beliefs*, *desires*, and *intentions*); and *action selection* (which determines the action to perform by considering current *intentions*). The *option generation* process decides how to achieve intentions, and, hence, is the one responsible for the agent’s *means-end* reasoning. On the other hand, *deliberation* is performed by the *filter* process. This process not only adopts new intentions in order to facilitate existing ones or to discover new opportunities. It can also drop intentions<sup>1</sup> that are not workable, and can retain intentions that have not yet been achieved [Weiss, 1999].

### 2.4.1.2 Conceptual Systems

By analyzing the main processes derived from the relation affect-cognition and affect-behavior<sup>2</sup> presented in Section 2.2, and by observing Table 2.1, it is easily observable that the relation of affect with the agent’s practical reasoning is mainly established with the processes of practical reasoning related to *perception* (or *belief revision*) and *deliberation*. In order to summarize the processes and attributes of Table 2.1, we propose the conceptual systems of Figure 2.1, which are based on the theories analyzed in Section 2.2. This figure

---

<sup>1</sup>Intentions that are either no longer achievable or that diminish the agent’s expected gains.

<sup>2</sup>Hereinafter we will use “affect-cognition” and “affect-behavior” to refer to “emotion-cognition” and “emotion-behavior”. It is a more general way of describing the relations between the affective state and cognition or behavior.

## 2.4. Affective Agents with Practical Reasoning

---

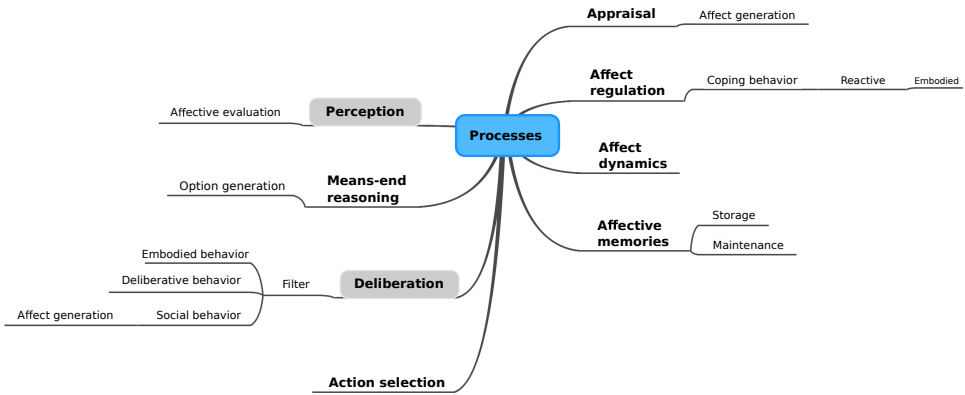
also shows the above-mentioned relation of affect with the agent practical reasoning.

Figure 2.1(a) left shows the processes of practical reasoning. The processes of *perception* and *deliberation* are highlighted because they are the processes of practical reasoning that have a relation to affective processes. Their sub-processes show how this relation is established. Figure 2.1(a) right shows the main affective processes. In addition to the processes of practical reasoning, Figure 2.1(a) also includes:

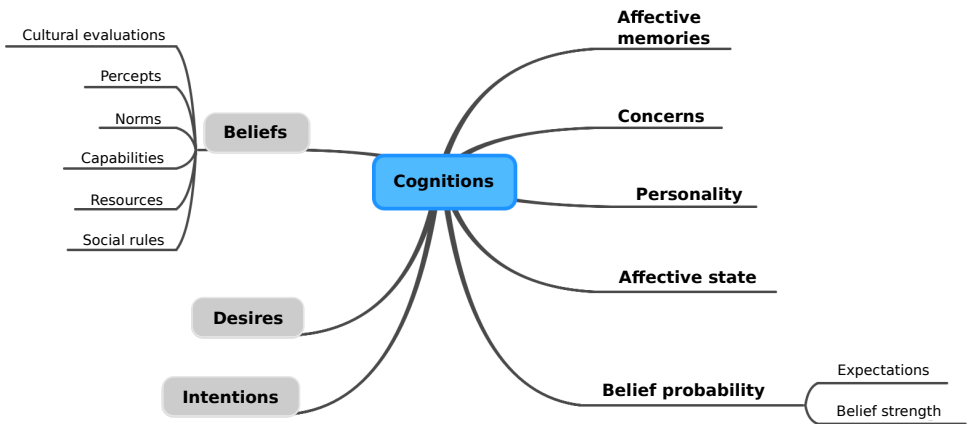
- *Affect generation*: it can take place as the result of social behaviors or as the result of the appraisal process. The affective state is modified according to the new stimuli.
- *Affective evaluation*: it can be part of the belief revision process (which includes perception) when an affective evaluation is performed regarding the content and the strength of new beliefs.
- *Embodied behavior*: an embodied behavior can be generated in the filter process, and it can be part of the affect regulation process (e.g., reactive facial expressions or body gestures).
- *Deliberative and social behavior*: the behavior associated to the filter process can also be either the result of a deliberation process, or it can be part of a social interaction.
- *Storage of affective memories*: it determines what affectively relevant information should be stored.
- *Maintenance of affective memories*: it determines how long affective memories are stored and how they change over time.
- *Coping behavior*: in general, the process of affect regulation leads to processes that are related to coping behavior. It includes processes for producing reactive behaviors, which in turn can include embodied behaviors.

## 2. STATE OF THE ART

---



(a) Processes



(b) Attributes

Figure 2.1: Conceptual systems for a BDI affective agent.

## 2.4. Affective Agents with Practical Reasoning

---

Figure 2.1(b) left shows a conceptual system that links the BDI components that are related to the *perception* and *deliberative* processes (i.e., beliefs, desires, and intentions) with other affect-related attributes whenever possible. Figure 2.1(b) right also includes new attributes that a BDI affective agent should have according to the main affect-related cognitive theories of Section 2.2.

Figure 2.1(b) shows that the attributes “affective memories”, “concerns”, “affective state” as well as the “belief probability” should be included in the definition of a BDI affective agent, in addition to the traditional BDI components beliefs, desires, and intentions. Based on Section 2.3, we also state that the agent should include an attribute for representing the agent “personality”. “Affective Memories” should be used by the related process of memory maintenance. “Concerns” are key for the appraisal process, and the “affective state” is the result of the affect generation process. On the other hand, “belief probability” helps in the representation of the agent’s expectations<sup>1</sup> and/or in the representation of the strength of beliefs (determined in the affective evaluation process). The probabilities of beliefs are also used by the appraisal process for the generation of, for example, prospective emotions [Ortony *et al.*, 1988]. In order to make a more reliable appraisal process, it would be ideal for the agent to be able to count on attributes like “cultural evaluations”, “percepts”, “norms”, “capabilities”, “resources”, or “social rules”. Since the use of these attributes may vary depending on the appraisal theory used, we haven’t included them as part of the “main” agent attributes but rather as special representations of the agent’s beliefs. Nevertheless, specific implementations can represent these attributes as additional agent attributes instead of being part of the agent’s beliefs.

### 2.4.2 Use of Core Processes and Attributes in Computational Approaches

In this section, we make a review of recent computational approaches of intelligent affective agents that have included, to a greater or lesser extent,

---

<sup>1</sup>For the representation of expectations on the basis of beliefs and their probabilities, the definition of a temporal system is also required.

## 2. STATE OF THE ART

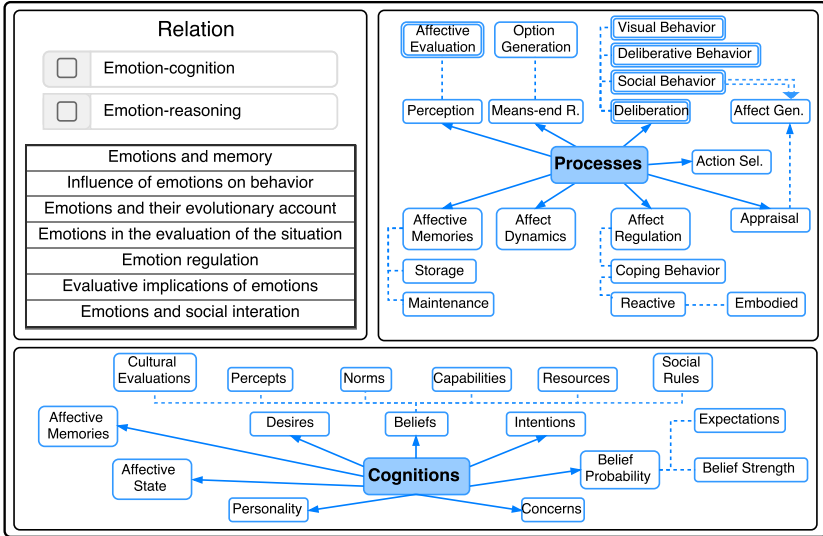


Figure 2.2: Theoretical framework for intelligent affective agents.

the main concepts and processes identified in Section 2.4.1. These concepts and processes have been integrated in the theoretical framework depicted in Figure 2.2. The upper left side of Figure 2.2 shows the possible affect-cognition and affect-reasoning relations that can be present in a computational approach. The upper right side shows specific processes, and the bottom shows specific cognitions that can be addressed by a computational approach. The processes and relations depicted in rectangles with double lines are those that may link processes of practical reasoning and affect-related processes. We consider that the extent to which affective computational approaches model these double line components indicates the extent to which they integrate practical reasoning and affect-related issues. This will help to determine the concepts and processes that are generally included in computational approaches according to their kind and domain of application. We will evaluate the selected computational approaches on the basis of the extent to which these approaches cover the main concepts and processes that we propose should be considered when designing a BDI affective agent. This evaluation is performed with the support of our theoretical framework.



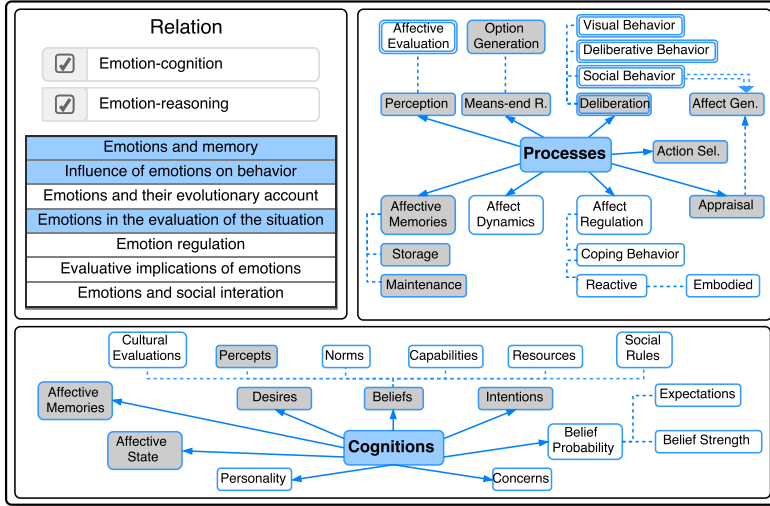


Figure 2.3: Processes and cognitions of the LIDA architecture.

### 2.4.2.1 Affect and Memory

An example that addresses the relation of affect with cognition is the work proposed in [Franklin *et al.*, 2014]. In this work the authors propose LIDA, which is composed of an agent architecture ( $LIDA_I$ ) and an agent model ( $LIDA_C$ ), for cognition, emotion, and learning that is inspired in cognitive science and empirically grounded in cognitive neuroscience. According to the authors, it is “a plausible candidate for a unified theoretical framework of cognition”. Autonomous agents in  $LIDA_I$  must continually sense the environment, interpret perceptions, and then act, where the actions must be appropriate to achieve the agent’s goals. LIDA mainly combines action selection, motivation via emotions, an attention mechanism, and learning.

Figure 2.3 shows the processes related to  $LIDA_C$  (a conceptual model where the emotion components still require computational testing), which include *Perception*, *Means-end Reasoning*, *Deliberation*, *Action Selection*, *Appraisal*, and *Affective Memories*. In LIDA, both the affect-cognition and affect-reasoning relations are considered. The first relation takes place through the influence of emotion on learning where, up to a point, there is a direct correlation between the levels of arousal that are involved with current emotions and the encoding

## 2. STATE OF THE ART

---

of information in memory. Beyond that point, arousal becomes a hindrance to learning. The second relation happens when the likelihood for actions to be selected depends on current emotions. In LIDA, “schemes” of possible actions have a context and a result as well as an activation value. Emotions increase or decrease the activation value of those actions.

In LIDA<sub>C</sub>, *Affective Memories* are stored and maintained. Moreover, memory is categorized according to the type, function, and duration of the structures learned. This categorization includes the following systems: sensory, perceptual, spatial, episodic, attentional, procedural, and sensory-motor. Specifically, in the perceptual system, a perceptual associative memory (PAM) stores, “nodes” containing an emotion and its intensity. An appraisal process links these nodes with “node/links”, which include the perceptual structure that represents the event that caused the emotion. Thus, this appraisal process includes an *Affect Generation* process.

For the perception process, there is an “understanding phase” in the LIDA cognitive cycle, where the representation of the agent relationship with its environment is also stored in the PAM. In addition, in an “action and learning” phase and as part of the deliberation process, templates (or schemes) of possible actions are instantiated with the current situation by evaluating the intersection of the context of those templates with the situation. Then, an action selection mechanism chooses a single action from the instantiated templates. The means-end reasoning process in LIDA is performed through procedural learning (where new actions and sequences of actions for accomplishing new tasks are learned). A procedural memory stores the resulting executing behaviors. This procedural learning corresponds to the *Option Generation* process.

There is not an explicit reference for the processes or cognitions that are not highlighted in Figure 2.3. Nevertheless, in [Franklin *et al.*, 2014] is argued that some of these processes or cognitions can be represented by PAM nodes and their activations.

### 2.4.2.2 Affect Influence on Behavior

An example of a relevant approach where affect influences deliberative and embodied behavior is WASABI [Becker-Asano, 2008; Becker-Asano & Wachsmuth, 2010]. WASABI (**A**ffect **S**imulation for **A**gents with **B**elievable **I**nteractivity) is an Affect simulation architecture that combines cognitive reasoning capabilities with simulated embodiment. This work was used in a Skip-Bo card game scenario for simulating emotion capabilities in a virtual player called MAX. Three agents interact to control the overall functioning of the virtual player: the *Emotion-Agent*, the *BDI-Agent*, and a *Visualization-Agent*. WASABI is BDI-based, and it models intentional actions in the form of plans, which perform the *Means-end Reasoning* and *Option Generation* processes. Figure 2.4 shows the relation between WASABI and the main processes and cognitions of our theoretical framework.

This architecture makes the distinction between primary and secondary emotions following the theories by [Damásio, 1994], [LeDoux, 1998] and [Ortony *et al.*, 2005]. Thus, emotions that are considered to be “primary” emotions have a direct relation with expressive capabilities like facial expressions, body postures, or voice inflection. They are considered to be infant-like emotions such as anger, happiness, or surprise. “Secondary” emotions arise as the result of reasoning about current events, taking into account expectations and past experiences; they are “prospect-based” and can be expressed verbally (e.g., hope or relief). In this approach, the *affective state* or *mood* is modeled as a background state whose value moves on a bipolar scale of positive versus negative. This value moves toward a neutral value more slowly than emotions. Each primary emotion is located in a PAD space (which stands for Pleasure, Arousal, and Dominance according to [Mehrabian, 1996b]) following the values from [Russell & Mehrabian, 1977]. The architecture of WASABI has six basic components: the conscious appraisal, the non-conscious appraisal, memory, deliberation, emotion dynamics, and the PAD space. The conscious and non-conscious appraisal components are in charge of the *Appraisal* of the stimuli from the environment through the *Perception* process. The appraisal process is based on Ortony’s and Scherer’s appraisal theories [Ortony *et al.*, 1988;

## 2. STATE OF THE ART

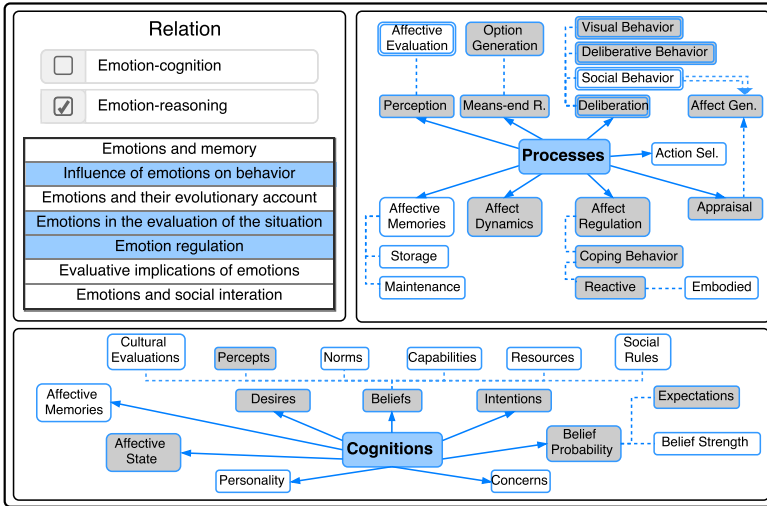


Figure 2.4: Processes and cognitions of the WASABI architecture.

Scherer, 2001]. As a consequence, secondary emotions and emotional impulses are generated by the “conscious” and “non-conscious” appraisal components, respectively, in an *Affect Generation* process. The way the conscious appraisal obtains secondary emotions is by evaluating the goal-conduciveness of an event. Then the memory is updated, and *Expectations* are generated.

The level of dominance is derived from this process, and is used by the “PAD space”. Mood is derived from the *Affect Dynamics* component as well as the Pleasure and Arousal values of the PAD space. Finally, primary emotions, which lead to involuntary behavior, are elicited by the PAD space with a specific intensity. On the other hand, the deliberative actions are generated by the *Deliberation* component. To do this, this component takes the resulting “aware emotions” obtained by the PAD space. This PAD space acts as an awareness filter. This process is done to keep the mood-congruency of the primary and secondary emotions. This congruence is based on the idea that, with a positive mood, we are less likely to experience negative emotions and vice versa. Unlike primary emotions, secondary emotions have a “lifetime” function and a “decay” function, so their intensity decreases over time until the zero base intensity is reached. The “aware emotions” (with positive awareness

## 2.4. Affective Agents with Practical Reasoning

---

likelihoods) may result in different coping behaviors after a deliberative process and a reappraisal process.

The approach in [Hoey *et al.*, 2016] is another example of the influence of emotions on agents' behavior, specifically on social behavior. The authors proposed a method where the behavior of human-interactive agents follows the "affect control principle". According to this principle, affective consistency is maintained in such a way that agent action choices seek to create situations that are consistent with those affective sentiments that are culturally shared or seek to suppress situations that are inconsistent with those affective sentiments that are culturally shared. To this end, the authors propose the BayesAct model, which is a partially observable Markov decision process. This model is a probabilistic decision theoretic generalization of the Affect Control Theory (ACT) [Heise, 2007], which is a mathematical formalization of the interactions between cultural representations, the kind of interactants in social situations, and affective experience. In this approach, the agent not only keeps a record of its current affective state (referred to as sentiments) and what "should be" its affective state, but it also records these same states in relation to the user. These records are called "identities". The BayesAct model is also able to learn about people's identities. Identities are modeled over three dimensions (the EPA dimensions): evaluation, which indicates how good or bad something is; potency, which indicates how strong or weak it is; and activity, which indicates how active or passive it is. These dimensions correspond to the dimensions of Mehrabian's PAD space [Mehrabian, 1996b].

Figure 2.5 shows the relation between BayesAct and the main processes and cognitions presented in our theoretical framework. In BayesAct, *Perception* gives rise to observations, which include information about the environment and about the user. The *Deliberation* process is performed by the BayesAct control process, which, in turn, performs the *Means-end Reasoning* and *Option Generation* processes by trying to keep affective consistency. The agent behaviors always have a social and communicative purpose. The *Appraisal* process can be considered to be a sense-making process where the interpretation of observed events is performed by emphasizing the cultural aspects of these interpretations and relying on the role of the language in this process. In this

## 2. STATE OF THE ART

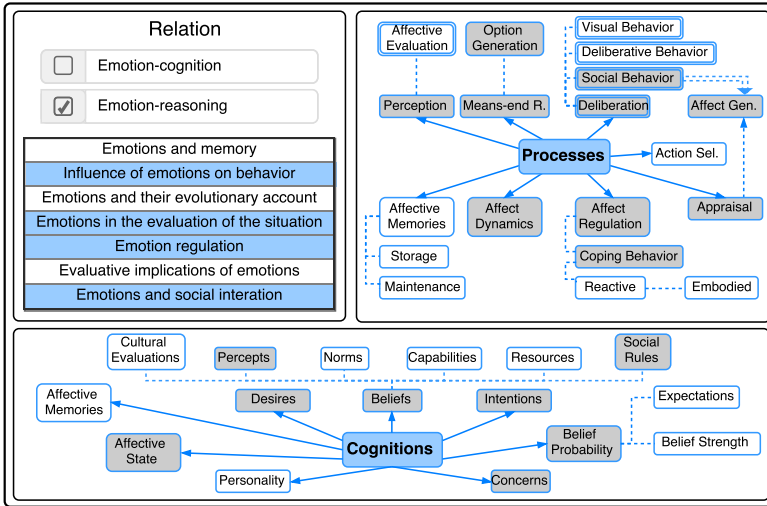


Figure 2.5: Processes and cognitions of the BayesAct model.

sense-making process, linguistic categories are organized by the EPA dimensions of the affective space. The whole process of BayesAct determines the *Affect Dynamic* over a continuous dimensional space, which, in turn, behaves as an *Affect Regulation* process, where deliberation creates behaviors that are oriented towards minimizing affective inconsistency. With regard to agent’s cognitions, in this approach, there isn’t an explicit representation for *Beliefs*, *Desires*, or *Intentions*, but they are somehow present in the model. Moreover, the knowledge (or beliefs) about an agent’s own sentiments and about the sentiments of other agent is represented as a probability distribution, so we consider that *Belief Probability* is represented in this approach. Besides, the model performs heuristic prescriptions for the actions of agents that interact with each other, which can be considered as a kind of *Social Norm*. The model also contains a utility function that is used in the action-choices, which can be considered as the agent’s *Concerns*.

### 2.4.2.3 Emotions and their Evolutionary Account

The work proposed in [Breazeal *et al.*, 2005] is an example that shows the role of imitation of emotions in social contexts (specifically in the context of an

## 2.4. Affective Agents with Practical Reasoning

---

infant and a caregiver), as part of an evolutionary adaption to improve social interactions. Imitation facilitates forms of social understanding like empathy or social referencing. As part of the imitation process, infants recognize structural congruence between themselves and the adult model, what serves as feedback for learning connections between the sensation of an action and its visual perception. Social referencing is a form of learning which is socially guided. In other words, it is learning in which a person formulates his or her interpretations of a given event and determines how to interact with it by using other people's interpretations. Unlike the approaches described above, this approach models two global issues: the imitation process and the process of social referencing. A robot architecture is proposed for implementing the imitation process in an infant, and a computational model is proposed for social referencing. In the imitation architecture, the "Perception system" extracts information from sensory input, which mainly consists of information that is related to the other person's facial expressions or movements. An "Action System" arbitrates the robot's behavior by selecting the appropriate behavior when required. Behaviors are represented as "action-tuples". The decision of what action-tuple will be executed is made by "action-groups", which are groups of action-tuples. The actions selected for execution are goal-directed, where goals mainly include the facial expressions the robot must imitate. Finally a "Motor System" is responsible for carrying out the selected action.

Since the influence of emotions is considered in the model of social referencing (which is called emotional referencing), we focus on this model to find out the relation between this approach and the main processes and cognitions proposed in our approach (see Figure 2.6). As part of this computational model, two models are proposed: the model of basic emotions, and the model of shared attention. In the model of basic emotions the *Appraisal* process simply tags *Percepts* and internal states with affective information (e.g., valence or novelty) by following the Damásio's "somatic marker" theory [Damásio, 1994]. The association between those percepts with affective information is learned by the infant robot through previous experiences of these affective states by mimicking the human facial expressions<sup>1</sup>. An affective state is generated in an *Affect*

---

<sup>1</sup>This assumption is made on the basis that some experiments with humans have found

## 2. STATE OF THE ART

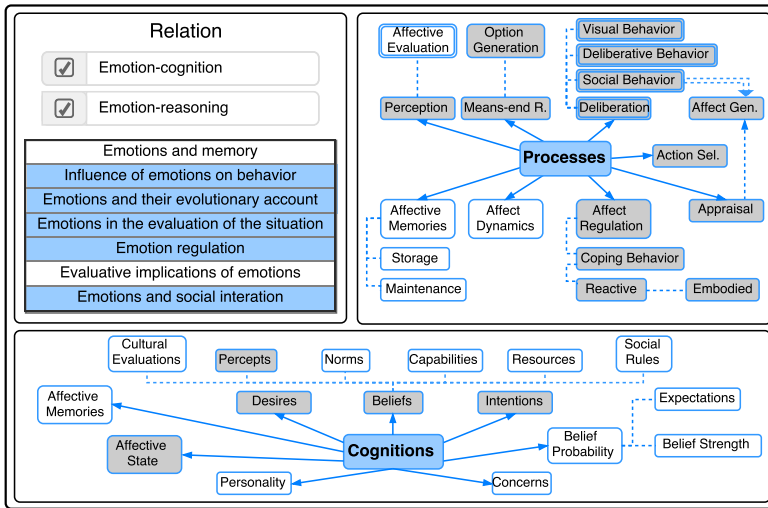


Figure 2.6: Processes and cognitions of the emotion model of the infant robot.

*Generation* process, which is represented through a set of widely accepted “basic emotions” [Ekman & Oster, 1982]. This change in the affective state may produce expressive or behavioral responses to cope with it. This can be seen as a kind of *Affect Regulation*, where the *Coping Behavior* is oriented to establish a desired relation between the robot and the environment. The model of shared attention mainly keeps track of the referential focus of both the infant robot and the human (e.g., an object they are looking at). The information of shared attention is used by the robot to associate the appraisal communicated by the human with the referential focus (i.e., the last object whose attention was shared by both the robot and the human). The model of social referencing also includes a *Perception* process (whereby the state of the others and of the world is perceived) as well as a *Deliberation* and *Option Generation* processes (in order to generate *Embodied*, *Deliberative*, or *Social* behaviors). The robot’s main goal is to refine its understanding of the world by using affective information.

---

that showing a facial expression that is associated with an emotion produces that emotion [Strack *et al.*, 1988].



### 2.4.2.4 Affect in the Evaluation of the Situation

This aspect is directly related to the process of *Appraisal*, which has probably been the most addressed issue by computational approaches. In fact, every approach that has been used as example in this chapter provides a description (more or less elaborated) for the appraisal process. We use as example the approach proposed in [Steunebrink *et al.*, 2012], which purely addresses the appraisal process by offering a formalization of the conditions that trigger emotions for BDI agents. The appraisal theory that underlies this formalization is the OCC model of emotions [Ortony *et al.*, 1988]. The formal model that results from this formalization is an extension of the KARO framework [Meyer, 2006; Meyer *et al.*, 1999]. This is an agent specification framework that can be considered a BDI logic based on dynamic logic. Through this framework, the agents' informational and motivational attitudes can be specified (e.g., *Beliefs*, *Desires*, *Goals*, and commitments). The platform also allows the agent behavior to be described.

This approach assumes the existence of a *Perception* process where, according to the OCC model, *Percepts* are grouped as: consequences of events, consequences of actions of agents, or percepts of objects. After the *Perception* process, the *Appraisal* process may take place, where percepts are evaluated according to criteria that are related to appraisal variables that are hierarchically structured in the OCC model. Some of these variables include desirability/un-desirability of the consequences of an event, praiseworthiness/blameworthiness of an action of an agent, and appealingness/unappealingness of the aspect of an object. These variables, along with the structures and the logic provided by the KARO framework, help to derive all OCC emotion types through an *Affect Generation* process. The processes and cognitions described in this extension of the KARO framework that are consistent with those proposed in our theoretical framework are highlighted in Figure 2.7.

### 2.4.2.5 Affect Regulation

In order to illustrate how a computational approach has addressed the processes whereby individuals perform (reactive or more planned) behaviors in order to

## 2. STATE OF THE ART

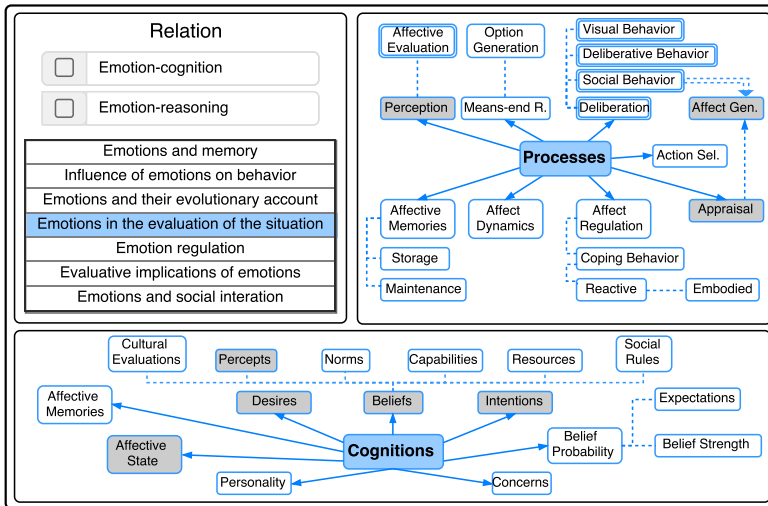


Figure 2.7: Processes and cognitions of the formalization of OCC appraisal.

manage situations generated by the emotions' elicitation, we describe the model proposed in [Adam & Longin, 2007; Longin, 2007]. The authors extended a previous work where the process of appraisal was formalized in a BDI framework. Their proposal starts from the concept of coping strategy of [Lazarus & Folkman, 1984], which is defined as the behavior that is oriented to the minimization or control of stressful situations. The logical framework of the approach is built on the basis of concepts such as formulas (which are the agent's *Beliefs*), actions, choices or preferences (which can be considered the agent's *Concerns*), *Desires*, *Expectations*, *Belief Probability*, and *Intentions*. These concepts lead to the formalization of an appraisal process where a set of emotions is derived, specifically, negative emotions, which are considered mental states, and whose conjunction can be considered as the *Affective State* of the agent. These concepts are also used in the formalization of emotion-focused and problem-focused coping strategies. According to the definition of coping actions schemes for these coping strategies, a coping action would take place when a specific emotion is experienced and the context of the situation is appropriate for applying the coping strategy. A formalization is also made in order to determine what coping action will be executed over a set of applicable coping

## 2.4. Affective Agents with Practical Reasoning

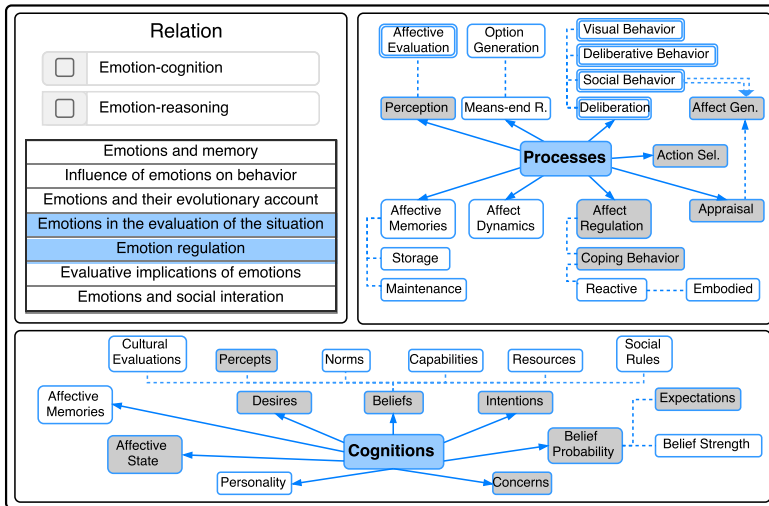


Figure 2.8: Processes and cognitions of the model of coping strategies.

actions, which can be seen as a mechanism of *Action Selection*. According to the authors' description, the selection of a coping action involves the processes of *Perception*, *Appraisal*, *Affect Generation*, selection of applicable coping strategies, selection of the preferred coping strategy, and, finally, execution of the corresponding coping action. These processes and concepts are highlighted in Figure 2.8.

### 2.4.2.6 Evaluative Implications of Affect

The approach in [Pimentel & Cravo, 2005], describes how the *Affective State* can influence the maintenance of *Beliefs*. It proposes an affective Belief Revision system whose goal is to keep the consistency of beliefs through tasks of belief revision that are influenced by the affective state. It is a logic-based system that uses the current affective state (or mood) in order to set a preference between new and old information. This preference is determined on the basis of the influence of anticipated emotions in the support or prevention of the acceptance or rejection of information [Frijda *et al.*, 2000]: for example, the “discouragement” of the process of acceptance or the abandonment of a belief, if this process or belief leads to the anticipation of a negative emotion; and

## 2. STATE OF THE ART

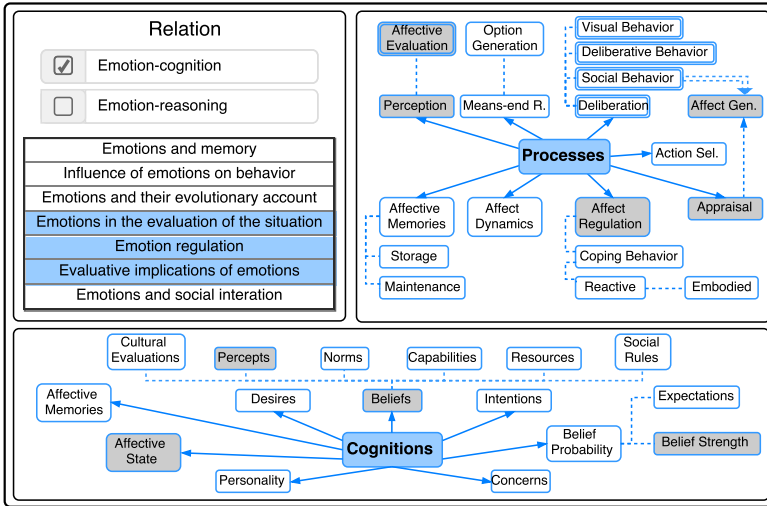


Figure 2.9: Processes and cognitions of the Affective Belief Revision system.

the “encouragement” of the process if the process leads to the anticipation of a positive emotion. This can be also seen as a process of *Affect Regulation*. The influence of affect on the preference between new and old information can also be described by the tendency of individuals to rely on their current beliefs when experiencing a positive mood, or to search for new information when experiencing a negative mood. Also, an ordering is endowed to beliefs regarding their resistance to change (i.e., the strength of beliefs). The strength of beliefs can be determined according to either the anticipation of emotions, or according to the credibility/authority of the source of the belief.

In order to model how emotion anticipation influences *Belief Strength*, a system based on OCC is proposed. It consists of a subsystem for *Appraisal* and a subsystem for calculating the intensity of emotions through an *Affect Generation* process. In the *Perception* process, *Percepts* related to the environment or to communications from other agents are considered. Figure 2.9 highlights the processes and concepts addressed in this approach.

### 2.4.2.7 Affect and Social Interaction

Several computational approaches have focused on the influence of emotions on the social aspect of agents in the same way as psychological theories have done. This study is useful in agents that must communicate with other agents or humans in a way similar to the way humans do. Some of the approaches used as examples in previous sections (such as the BayesAct model of Section 2.4.2.2 and the infant robot of Section 2.4.2.3) can also be used in the present section as examples. Nevertheless we have selected the computational model of empathy described in [Boukricha *et al.*, 2013] because it specifically models a social phenomenon where emotions play an important role.

In [Boukricha *et al.*, 2013], the authors start from the idea that empathic behavior can be modulated by factors such as the agent’s affective state, personality, or level of relationship with the agent’s interactant. In this approach, empathy is modeled in a virtual human called EMMA. Two modules are integrated in the EMMA cognitive architecture (an emotion simulation module and a BDI module). The emotion simulation module is based on the approach presented in [Becker-Asano & Wachsmuth, 2010], which was previously discussed in Section 2.4.2.2. We highlight the same processes and cognitions of Section 2.4.2.2 in the present example (see Figure 2.10). The BDI module follows a traditional BDI architecture, where *Perception*, *Deliberation*, and *Means-end Reasoning* are integrated [Leßmann *et al.*, 2006]. This BDI module also has other characteristics. For example, it has a reactive component that connects perception and action in order to generate immediate responses; the process of deliberation also decides what action to take next by composing interaction moves. These interaction moves are actions (e.g., communicative or manipulative like “connect” or “rotate”) that are combined to achieve a desired state of the world.

The empathy mechanism proposed in [Boukricha *et al.*, 2013] starts from a perceived facial expression and imitates this facial expression through an internal imitation process. Then, an emphatic emotion is elicited, which is followed by an emphatic modulation process. The emphatic modulation process modulates the emphatic emotion taking into account factors like “liking” or

## 2. STATE OF THE ART

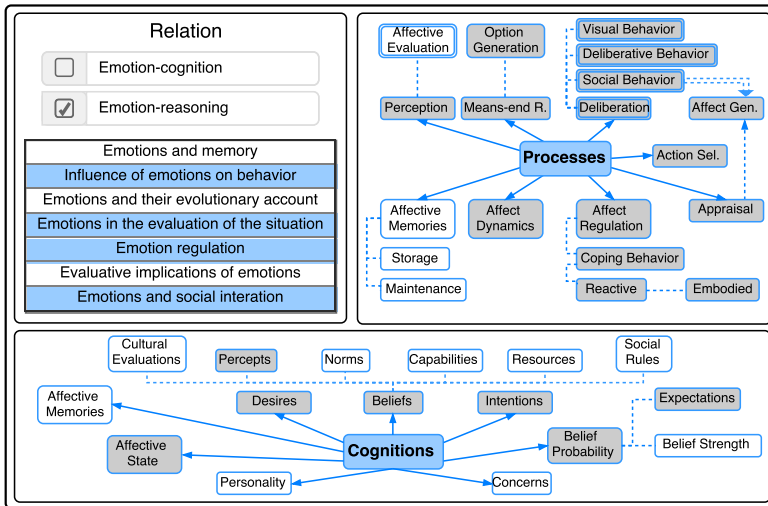


Figure 2.10: Processes and cognitions of the computational model of empathy.

“familiarity”. When experiencing the emphatic emotion, *Social Behaviors* like facial expressions, body gestures, or specific dialogs can be triggered in the virtual character EMMA.

## 2.5 General Approaches

Several authors have proposed specialized mechanisms to incorporate affective components into intelligent agents. Usually these components represent an appraisal derivation model, an affect derivation model, and a behavior consequent model [Marsella *et al.*, 2010]. These models have been used in several computational models like EMotion and Adaption (EMA) [Marsella & Gratch, 2009], Fuzzy Logic Adaptive Model of Emotions (FLAME) [El-Nasr *et al.*, 2000], MEA [Battaglini *et al.*, 2013], or PEACTION [Marinier *et al.*, 2009] (see [Reisenzein *et al.*, 2013] or [Rodríguez & Ramos, 2014] for a detailed review of affective computational models). According to these models, relatively general affective architectures have also been proposed. Two relevant examples are the Methodology and Architecture for Modeling state and traIt effects on Decision-Making (MAMID) [Hudlicka, 2004, 2008] and Fearnot AffecTive

Mind Architecture (FAtiMA) Modular [Aylett & Paiva, 2012; Dias *et al.*, 2014].

MAMID implements a processing cycle that includes sensing, thinking, feeling and acting. This processing is performed through different modules whose computations are supported by long-term memories. The different types and structures of these memories differentiate agents regarding their histories, sensitivities, and responsiveness. MAMID models the processes of emotion generation, emotions' effects on processing (i.e., on allocation of processing resources, on the selection of actions or goals, and on the speed and capacity of its different modules), and specially, it models emotions' effects on attention, perception, and cognition. It provides flexibility to the processes of emotion generation and emotions' effects in such a way that it allows to parametrize the processes and structures of its different modules. Specifically, these parameters are calculated through linear combination of the factors that influence each parameter. On the other hand, FAtiMA Modular [Aylett & Paiva, 2012; Dias *et al.*, 2014], is a BDI architecture that uses emotions and personality to influence agents' behavior. FAtiMA is a modular approach that has a core structure with a general processing cycle (and generic functions) for the agent. These generic functions are performed through specialized components (e.g., a reactive component, a deliberative component, a motivational component, a cultural component, etc.). New specific components can also be implemented with additional functionalities, or different psychological theories.

Considering our theoretical framework, both MAMID and FAtiMA Modular (and its specialized components) include structures for modeling both emotion-cognition and emotion-reasoning relations. They also model most of the processes and cognitions of our theoretical framework. Nevertheless, although MAMID is flexible regarding the level of parametrization, it is not so regarding the way these parameters influence cognition and reasoning (always a linear combination of the parameters). Besides, affect generation is limited to two appraisal levels (valence and one of the basic emotions fear/anxiety, anger, sadness, joy). FAtiMA Modular, in turn, can contain all the cognitions and processes proposed in the theoretical framework, regarding its modular nature.

### 2.6 Conclusions

Although several reviews of computational approaches that model affective processes on agents have been performed, few of these reviews have established in detail the links between these approaches and the philosophical and psychological theories that motivate them. These links are often discussed in computational approaches that model a specific issue that is related to affect, but these links are rarely addressed in a generic and abstract way. By addressing these links and concepts in a generic way, generic computational approaches of affective agents with practical reasoning can be created. These generic computational approaches can provide basic structures that allow the implementation of different specialized affect-related theories. In this chapter, practical reasoning, affect-related aspects and their integration have been discussed from a theoretical point of view. This analysis sheds light on the concepts and processes that a generic computational approach should model in order to build affective agents with practical reasoning that live in an agent community. These concepts and processes have been integrated in a theoretical framework that fulfills the Goal G1 of this thesis, which has been described in the introduction of this chapter. We have also analyzed how some representative examples of computational approaches have addressed these concepts and processes. These examples are very different regarding the domain of application for which they were built. However they jointly cover the main processes and concepts addressed by inspiring philosophical and psychological theories.

The method used in this chapter to analyze the coverage of computational approaches regarding the processes and concepts proposed has some advantages. First, it can be extrapolated to evaluate how comprehensive and well grounded an approach is regarding the processes and concepts we propose should be modeled in an affective agent with practical reasoning. Second, the incremental development of mechanisms and tools is facilitated since it becomes easier to identify functionalities to be added on the basis of processes that are not modeled. Third, this method makes it possible to determine to what extent practical reasoning and affect-related issues are modeled and to what



Table 2.2: Summary of characteristics of the computational approaches.

Approach	Kind	App. Domain	No. processes		Integration PR - Affect
			PR	Affect-Rel.	
LIDA	Architecture and model	Cognition, and learning	5	5	0
WASABI	Architecture	Cognition and simulation	6	6	3
<i>BayesAct</i>	Model	Social context, cultural aspects	5	5	2
Infant robot	Architecture and model	Social context, evolutionary adaption	8	6	4
Coping strategies	Formalization	Coping strategies	2	4	0
Affective Belief Revision	Formalization	Beliefs revision	1	4	1
OCC appraisal	Formalization	Appraisal process	1	2	0
Model of empathy	Model	Social context, empathy	8	7	4

App. Domain: application domain, PR: practical reasoning, Affect-rel.: affect-related.

## 2. STATE OF THE ART

---

extent they both are integrated. Specifically, some conclusions have also been extracted from the approaches analyzed. Table 2.2 shows a summary of their characteristics, including what they model, application domain, number of processes modeled (practical reasoning processes and affect-related processes)<sup>1</sup>, and level of integration of practical reasoning and affect-related issues (see Figure 2.2).

Table 2.2 shows that practical reasoning has been modeled by most of the approaches that focus on specific affect-related processes. Only those approaches on which formalizations of a specific affective process are proposed (such as those related to “affective evaluation”, “affect regulation”, or “appraisal”) do not cover the basic processes of practical reasoning. Nevertheless, in order to properly validate these approaches, an agent with a reasoning component that is able to adopt these new features may be very necessary. The approaches analyzed here are representative and not fully comprehensive. However, the method used in this chapter allows also to conclude that, from the approaches analyzed, those that model social behavior (such as empathy or mimicry) are the ones that model the largest number of processes, which can be related to practical reasoning, affect, or their integration. As complex phenomena (e.g., mechanisms that involve their own models as well as models of other agents), it becomes evident that their modeling requires considering more processes than those needed by simpler phenomena.

It would be desirable for a generic computational approach to be able to cover, at least basically, the selection of core processes proposed in this chapter. This kind of generic approach is scarce. The theoretical framework that we have offered allows such generic computational approaches to be built, thus allowing these computational approaches to be philosophically and psychologically well grounded. However, many challenges still remain for these approaches to become applicable and practical. For example, it would be necessary to make a proper selection of default mechanisms for the processes proposed, in such a way that it allows the incremental improvement of these processes individually.

---

<sup>1</sup>We consider practical reasoning processes to be those shown in Figure 2.1(a) left except the “Affective evaluation” and the “Affect generation” processes. The rest are the affect-related processes.

Also, considering that every application domain does not require modeling all the processes proposed (e.g., if agents are not embodied, embodied behaviors wouldn't be required), a mechanism for easily “decoupling” processes that are not necessary may be required. Mechanisms for establishing different kinds of relationships between practical reasoning and affective processes are also desirable (so that agents can be designed to be more or less “emotive” or “rational”).

On the other hand, not only are generic platforms uncommon, but agent languages that are able to provide useful structures to program affective agents with practical reasoning are also uncommon. Current research is still focused on models, architectures, and the formalization of specific affective phenomena. However, to the best of our knowledge, no agent language has been proposed to allow the implementation of a wide range of kinds of affective agents with practical reasoning. The concepts (agent cognitions) that we have identified in this work constitute a good starting point to determine what this kind of agent language should include. By delimiting basic processes and concepts (which are related to either the appraisal process or to the affect influence on behavior), the effort of researchers and software designers will focus on using or adapting them to specific applications. Therefore, this is a new step towards standardization in emotion modeling, which is one of the current challenges in artificial intelligence [Hudlicka, 2014]. The theoretical framework that has been proposed in this Chapter will guide the design and formalization of a general purpose affective agent architecture that will be described in Chapters 3 and 4.



## Part III

# Affective Architecture



---

---

## GenIA<sup>3</sup>: a General purpose Intelligent Affective Agent Architecture

### 3.1 Introduction

Agents developers and computer scientists often face the challenge of starting from scratch when including affect in the representation of an agent and its behavior. Scientists use their “creativity” to define the way emotions are appraised, experienced, and their effects on behavior. But “the *direct* implementation of an emotion theory as a computer program is usually not the best way to go” and “this is a cumbersome way to proceed” [Reisenzein *et al.*, 2013]. In this thesis we make a further step toward the standardization process mentioned in Chapter 1. To this end, we start from the main processes and concepts of the generic framework proposed in Chapter 2 to propose *GenIA*<sup>3</sup>, a General-purpose Intelligent Affective Agent Architecture [Alfonso *et al.*, 2014b, 2016c, 2017]. Hence, *GenIA*<sup>3</sup> is based on some of the most widely accepted psychological and neurological theories presented in Chapter 1. *GenIA*<sup>3</sup> offers components to represent affective characteristics like personality, emotions and mood. It is built over a traditional BDI architecture, since the BDI agent architecture is a suitable and widely exploited alternative to model

### 3. GenIA<sup>3</sup>

---

intelligent agents. Being based on a BDI architecture, *GenIA*<sup>3</sup> follows theories of motivation and action generation, where the course of actions to be executed is decided according to the agent’s goals. In *GenIA*<sup>3</sup> the affective state also has a motivational function, guiding goal’s priorities and/or generating new intentions. The interaction among the components of *GenIA*<sup>3</sup> produces an agent behavior biased by the agent affective state.

By designing *GenIA*<sup>3</sup> we aim to achieve the Goal G2.1 of this thesis (see Chapter 1), by allowing to implement different psychological theories relative to individual differences, affect generation, affect dynamics, and affect influence on cognition and behavior, on the basis of a common conceptual system.

## 3.2 Processes of an Affective Agent Life Cycle

In this section we identify the main cognitive processes that are part of *GenIA*<sup>3</sup>. These processes are in line with theories of motivation and action generation. According to [Reisenzein *et al.*, 2013] “nearly all current theories of motivation and action generation are variants of a singly basic theory, the *belief-desire* theory of action”. These psychological theories of motivation have inspired computational emotion models, which have been used in cognitive agent architectures. In particular the belief-desire-intention (BDI) architecture [Bratman, 1987] is a practical and powerful conceptualization, widely accepted in the agent’s community which has been the base of numerous computational approaches.

The affective processes addressed in this thesis are in line with *cognitive-motivational* (or belief-desire) emotion theories, whose constituents or mental causes are both: beliefs and desires [Reisenzein *et al.*, 2013]. Thus, the motivational function of emotions is to guide goal’s priorities and/or to generate new goals. We start from the BDI model to identify the core cognitive processes of the life cycle of an affective agent, due to the dependence of emotions on beliefs and desires, and the constructs for building autonomous and goal-directed agents that BDI logics provide. In order to illustrate how each of these processes work, we will use the following example.

**Example 1** Consider an agent that imitates humans playing a card game



that involves gambling. The agent, then, needs mechanisms for making decisions about what move to make next, possibly biased by the current affective state of the agent, and mechanisms for how to emotionally react to what happens in the game<sup>1</sup>.

### 3.2.1 Processes in a BDI Agent

The BDI architecture has its foundations on the philosophical theory of “practical reasoning”, where the actions to be performed are decided according to particular goals. According to the BDI theory, rational agents are committed to their intentions, intending to do always what they believe will lead to their desires [Bratman, 1987]. Practical reasoning requires both: establishing what goals to achieve (deliberate) and how to achieve them (also called means-end reasoning), what can be summarized in four main processes performed consecutively [Weiss, 2013]:

**Belief revision function (brf).** It uses a perceptual input along with current *beliefs* in order to determine the agent’s new *beliefs*. As a result of this process, new *external events* (one per percept) may be generated.

**Option generation function (options).** It performs the tasks for *options* generation. These *options* (or desires) are generated on the basis of the agent’s current *beliefs*, *external events*, *internal events*, and *intentions*. These *options* represent the means whereby the agent can achieve its *intentions* (e.g., rising profits).

**Filter function (filter).** It determines what to do by generating the agent’s *intentions* (e.g., doubling the bet, standing, or hitting a card). To this end, a deliberation process is performed which considers previously-held *intentions*, current *beliefs*, and *options*.

**Action selection function (execute).** Contains the “action selection function”, thus it uses the current *intentions* to determine the next action to

---

<sup>1</sup>Literature argues for the influence of emotions on decisions in these kind of games [Bechara *et al.*, 2000; Camerer *et al.*, 2011].

be executed. The execution of actions can produce *internal events* that are related to, for example, the action failure or success. For example, in Example 1, if the action “rise bet” is executed, a possible reason for it to fail could be that the bet reached a top value.

#### 3.2.2 Agent’s Affective Processes

According to [Hudlicka, 2014] the agent’s core affective processes can be modeled across four emotions’ modalities of biological agents. These emotion modalities are the *behavioral/expressive* (which results in action-oriented characteristics and behavior); the *somatic/physiological* (which has a neurophysiological orientation); the *cognitive/interpretive* emotions (which are manifested as the result of the agent’s evaluation according to its goals, preferences, and the current situation); and the *experiential/subjective* (which results in an idiosyncratic and conscious experience of emotions). We focus on the *cognitive* emotion modality, and hence, on the corresponding processes for emotion generation and effects. We also focus on the *behavioral/expressive* emotion modality, specifically on the processes for emotion effects related to action selection.

The Goal G2.1 of this thesis is to propose an approach that allows to implement different theories relative to individual differences, affect generation, affect experience, and affect influence on cognition and behavior. Therefore, *GenIA*<sup>3</sup> includes the core processes that fulfill this requirement. First, individual differences are represented through personality traits that may influence the processes of the agent reasoning and affective cycles. Besides, affect generation is represented through the **appraisal** process, affect experience through the **affect generator** and **affect temporal dynamics** processes, affect influence on cognition through the **affective modulator of beliefs**, and affect influence on behavior is represented through the **coping** and **filter** processes. *GenIA*<sup>3</sup> also includes the processes of the traditional BDI agent architecture (see Section 3.2.1). Next, we describe these core affective processes.

Five core affective processes are included in *GenIA*<sup>3</sup>. These processes include:

**Appraisal.** Performs the evaluation of the current situation according to the current state of the world and the agent’s *concerns* (i.e., interests, needs, motivations, ideals, or standards). In *GenIA*<sup>3</sup> this process can be committed to any particular appraisal theory (e.g., [Smith & Lazarus, 1990], [Scherer, 2001], [Scherer *et al.*, 2001], or [Ortony *et al.*, 1988]), and several parameters can be used in order to perform this appraisal (e.g., the agent’s *beliefs*, *concerns*, *internal events*, *external events*, memories of *affectively relevant events*, *current options*, and *intentions*). A set of *appraisal variables* results from this evaluation. Consider, for example, that the concerns of the agent of Example 1 involve to win every single hand, and he loses the current hand. An appraisal variable, “desirability”, with a low value, may result from the evaluation of this event in the appraisal process.

New relevant events (according to their impact on the affective state of the agent), are also stored in the memory of *affectively relevant events* during the appraisal process. These affectively relevant events may be used in a future appraisal<sup>1</sup>. An example of an affectively relevant event could be to lose a risky hand where a high bet was made.

Although the appraisal process has several parameters, not all of them are necessarily used. For example, following the EMA model [Marsella & Gratch, 2009], the appraisal variable *desirability* is determined by assessing the value of a proposition. This may imply the use of the agent *concerns* and the agent *beliefs* in order to evaluate the agent concerns.

**Affect generator.** It is in charge of generating the agent *affective state* by using the current *affective state* and the *appraisal variables* generated by the appraisal process. The affect generator process can be committed to any related psychological theory and the agent *affective state* can be represented either as a set of emotion categories, appraisal variables, or mood dimensions. For example, when representing the *affective state*

---

<sup>1</sup>This possibility is allowed on the basis of theories that argue that past personally significant events (which are stored in the autobiographical memory), can have a significant impact on human life, shaping the perception of the upcoming tasks and modifying actual behavior [Conway, 1990; Schwarz, 2000; Selimbegović *et al.*, 2015].

through mood dimensions, the Pleasure-Arousal-Dominance or PAD model of mood [Mehrabian & Russell, 1974], or the Russell’s bipolar dimensions (hedonism and arousal<sup>1</sup>) [Russell, 2003] can be used. In Example 1, if the agent is very depressed and he wins a hand, he may feel happier, but probably not as happy as if his previous affective state were happy already. Thus the purpose of affect generator is determining the way the agent’s affective state changes.

**Coping.** The coping process is tightly linked to the agent *personality*, since this process determines whether some agent responses or some reactive behavior should be generated, and what should be these responses or reactive behavior. These agent responses can represent “response tendencies” in line with [Ortony, 2003], or can be oriented to take back the agent *affective state* to a desired state<sup>2</sup> (also called “coping strategies” [Marsella & Gratch, 2009]). Examples of possible reactive behaviors are facial expressions or body gestures that are involuntary and individual. The agent can also have coping strategies like “shift responsibility” (e.g., to think that he lost a hand because the dealer gave him bad cards), or “wishful thinking” (e.g., to think that he will win all subsequent hands). These coping strategies may imply the generation of new *intentions*, as well as the modification on the *beliefs* of the agent through the process affective modulator of beliefs (described below).

**Affects temporal dynamic.** It is in charge of determining the temporal variation of the *affective state*, specifically its duration and decay. These dynamics vary from one individual to another, in such a way that some *personality* traits can determine the way that these variations are produced. This process doesn’t depend of any other process and no other

---

<sup>1</sup>According to Russell’s definition of affective state: “A neurophysiological state that is consciously accessible as a simple, nonreflective feeling that is an integral blend of hedonic (pleasure–displeasure) and arousal (sleepy–activated) values” [Russell, 2003].

<sup>2</sup>The specification of a “desired state” depends on the assumptions or psychological theories used in a particular design. For example, according to [Gross & Thompson, 2011], a desired state is that where emotional distress is reduced, but according to [Lazarus & Folkman, 1984] it is a state where the negative emotional responses associated with stress are reduced.

### 3.2. Processes of an Affective Agent Life Cycle

---

process depends on it. It determines the duration of the affective state’s components as well as how their intensities decay over time. This process partially models the agent’s affect experience which is supported by theories of mood and affect such as [Frijda, 2007], [Verduyn *et al.*, 2013], and [Koval *et al.*, 2014].

**Affective modulator of beliefs.** This process is not executed as an independent process but as a subprocess of either the coping process or the **brf** process described in Section 3.2.1 that determines how *beliefs* are maintained. It follows the idea that the agent *affective state* contributes to the maintenance of *beliefs* [Ito *et al.*, 2010; Marsella & Gratch, 2003; Pimentel & Cravo, 2005] (e.g., a negative affective state induces an individual to question more his or her beliefs, making him or her more susceptible to accept new information). In order to understand the function of the **affective modulator of beliefs**, let’s consider the agent perceived self-efficacy in the card game: the belief related to “the conviction that one can successfully execute the behavior required to produce the outcome” [Bandura, 1977]. If he continuously loses several hands, his perceived self-efficacy may be affected (see [Pimentel & Cravo, 2005] and [Frijda *et al.*, 2000] for a more detailed description of affective beliefs revision).

Finally, although the **filter** process of Section 3.2.1 is not considered an affective process, in *GenIA*<sup>3</sup> it has been extended so that it can consider the current agent *affective state*, and some aspect of the agent’s *personality* as part of this deliberation process. Specifically, in relation to the agent’s personality, the individual characteristics that represent the extent to which the agent’s decisions are influenced by its *affective state* can be taken into account. We call it *rationality level*, which has been included to fulfill the Goal G2.5 of this thesis, by facilitating to set an equilibrium between the rational and the affective sides<sup>1</sup>. We include the affective state and personality aspects on the filter process, on the basis of theories that argue for the need of considering the

---

<sup>1</sup>Some tests like the Cognitive Reflexion Test (CRT) [Frederick, 2005] are used in psychology to determine this issue.

### 3. GenIA<sup>3</sup>

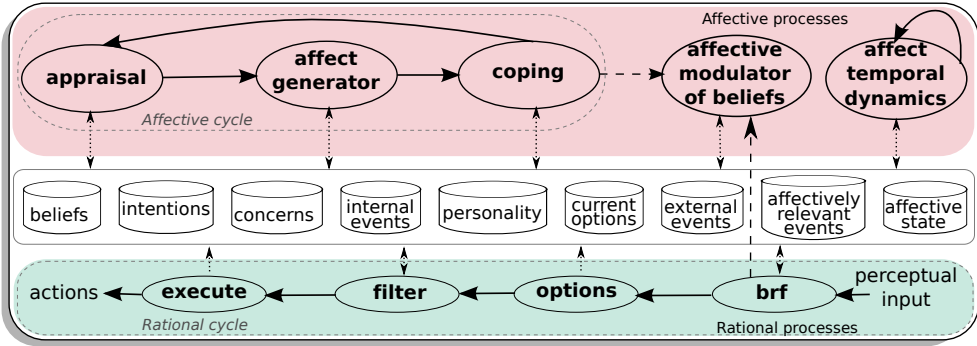


Figure 3.1: *GenIA*<sup>3</sup>. Sequences (solid line arrows), subprocess (dashed line arrows), and information exchange (dotted line arrows).

influence of emotions and individual differences for behaving either rationally or emotionally/intuitively/unconsciously in order to properly model human behavior [Kennedy, 2012]. Besides, experimental studies offer evidences that emotions drive deliberative decision-making [Camerer *et al.*, 2011].

The processes of Sections 3.2.1 and 3.2.2 are in line with the processes of the conceptual framework prosed in Chapter 2, which we concluded were essential according to our view of how a rational and affective agent should be built. Section 3.3 describes how these processes are integrated in *GenIA*<sup>3</sup>.

### 3.3 Integration of Affective and Rational BDI processes

In order to establish the relation and sequence of rational BDI and affective processes, in this section we describe *GenIA*<sup>3</sup> (a General-purpose Intelligent Affective Agent Architecture), which is shown in Figure 3.1. The sequence of the affective processes is independent from the sequence of the BDI processes. This means that, an agent that is built on top of this architecture may have, at least, one execution thread for each set of processes. Moreover there is an affective process which is “disconnected” from the others: the affect’s temporal dynamic, because its function of controlling variations on the intensity of the affective state doesn’t depend on other processes. The sequence of BDI processes (hereinafter *reasoning cycle*), includes the main processes of a BDI

### 3.3. Integration of Affective and Rational BDI processes

---

agent (bottom side of Figure 3.1). The architecture has two more cycles: one has only one process that is executed continuously (the *affect temporal dynamic*), and the other (hereinafter *affective cycle*) includes the rest of affective processes.

The affective cycle takes place as follows. When internal or external events are triggered, they are evaluated in the appraisal process. Other parameters, related to the cognitive information of the agent, are also used in the appraisal process (e.g., its concerns, beliefs, or the agent’s personality). The resulting appraisal variables are used by the affect generator, which updates the current affective state of the agent. By varying the agent’s affective state, the coping process checks if some action is required in order to either taking back the affective state to a desired state, or to perform some reactive action. Optionally, the coping process may require updating the agent’s beliefs, if this is one of the agent’s “coping strategies” [Ortony *et al.*, 1988]. This update is performed by the process affective modulator of beliefs. The appraisal process can be triggered again after the coping process, even when there isn’t an event to be processed. For example, if appraisal complies with Scherer’s theory (which is one of the more complex theories of appraisal), the appraisal is conceived as a multilevel sequential checking where a set of evaluation checks are performed in sequence [Scherer, 2001]. The Gratch and Marsella’s EMA model [Marsella & Gratch, 2009] also fits in this structure, because, in this model, affect is derived from a continuous cycle of appraisal, coping and re-appraisal. In general most appraisal theories can be represented in this way. The appraisal, affect generator, and coping processes are thus part of the appraisal-reappraisal cycle (also called *affective cycle*), which is represented in most appraisal theories<sup>1</sup>. The affect temporal dynamics, on the other hand, is not included in this affective cycle because it doesn’t depend on any other process and no other process depends on it. Thus, the affect temporal dynamics is controlled in an independent cycle.

Following appraisal theories, and fulfilling the goal of allowing to set an

---

<sup>1</sup>Appraisal-reappraisal is the term used in the Scherer’s appraisal theory [Scherer, 2001], which is considered one of the most complex and, hence, the representation of other appraisal theories could be easily done through it [Dias *et al.*, 2014].

### 3. GenIA<sup>3</sup>

---

equilibrium between the affective and rational sides of an agent (Goal G2.5 of this thesis), the implementation of the architecture allows to set a synchronization between the reasoning and affective cycles. This is described in Section 4.3.

On the other hand, although the BDI processes in the *GenIA*<sup>3</sup> architecture maintain their original functions, some of them include new functions. After the agent perceives its environment, it updates its beliefs in the *brf* process, which includes the affective modulator of beliefs in order to evaluate how the current affective state influences the content and the strength of the beliefs. Then, the available options are selected in order to determine the new course of actions in the *options* process. The *filter* process determines what to do by selecting the next intention, including the agent's affective state as a parameter. Finally the selected intention is executed in the *execute* process.

As shown in Figure 3.1, in *GenIA*<sup>3</sup> the interaction between affective and BDI processes is produced by using and updating the information related to the agent's cognition. These concepts become agent attributes that include: the agent's *beliefs* (i.e., the information the agent has about the environment, about himself, or about others); *intentions* (the actions the agent has committed to carry out); the agent's *concerns* (the agent's personal values, such as its interests, motivations, needs, standards, norms, or ideals); *internal events* (events that take place during the agent execution and that may generate changes on the agent state, for example, in the agent's affective state); *personality* (a representation of the agent's personality); *current options* (the options the agent has for facing the current situation, which is in line with the "desires" concept of the BDI model); *external events* (events perceived from the environment); *affectively relevant events* (those events that have produced important changes on the affective state of the agent); *affective state* (a representation of the affective state of the agent).

Some of these attributes may not vary along the whole life time of the agent, such as the agent's *personality* and its *concerns*. Other attributes, in turn, may change more frequently, for example, the *affectively relevant events*. The rest of the agent attributes can vary in every reasoning cycle,



### 3.4. Agent Attributes for BDI and Affective Processes

---

such as the agent's beliefs, internal events, intentions, current options, external events, and affective state. Table 3.1 shows the precise interaction of BDI and affective processes with these attributes. A more detailed description of these attributes can be found in Section 3.4.

## 3.4 Agent Attributes for BDI and Affective Processes

The main agent attributes that should be considered for modeling affective and BDI processes can be obtained by analyzing the inputs and outputs of these processes. By determining which are these agent attributes general agent languages, such as the one proposed in this work, could be built. This section explains in more detail these agent attributes, which were also introduced in Section 3.3, (see Table 3.1). It is also discussed why these attributes are considered as inputs or outputs for the processes of *GenIA*<sup>3</sup>. We begin this section by describing the representation of the agent attributes that are related to the BDI processes.

### 3.4.1 Representation for BDI Processes

The inputs and outputs of the *GenIA*<sup>3</sup>'s processes are shown in the last four rows of Table 3.1. As shown in Table 3.1, current `beliefs` are both inputs and outputs of the *brf* process, since its function is to revise beliefs. Likewise a representation of `external events` is created in this process, which are the events perceived from the environment. In order to determine the possible actions at a given time, the *options* process takes the current `external events`, `internal events`, `current intentions`, and the agent's `beliefs`. Then the current `options` are updated. The next action to be performed is selected starting from the current `options` in the *filter* process. This process considers current `beliefs` and the agent's current `affective state`, and updates `intentions`. Different agents have different propensities to be influenced by their affective state when making decisions. Therefore, the rationality level of the agent's personality can also be an input for the *filter* process. The *execute* process carries out the selected action, which uses the current

### 3. GenIA<sup>3</sup>

---

Table 3.1: Inputs and outputs for the processes of *GenIA*<sup>3</sup>.

Process	Inputs	Outputs
Appraisal	beliefs, concerns, internal events, external events, affectively relevant events, current options	affectively relevant events
Affect generator	affective state	affective state
Coping	personality (coping strategies), affective state	intentions, beliefs
Affective modulator of beliefs	affective state, beliefs	beliefs
Affect's temporal dynamic	affective state, personality	affective state
brf	beliefs	external events, beliefs
options	beliefs, external events, internal events, intentions	options (desires)
filter	personality (e.g. rationality level), current options, beliefs, intentions, affective state	intentions
execute	intentions	internal events

intentions and generates internal events like the intention's failure or success.

#### 3.4.2 Representation for Affective Processes

In *GenIA*<sup>3</sup>, the *appraisal* process is a part of a continuous appraisal-affect generation-coping cycle that may process an event. The goal of this configuration of the architecture is to make it flexible enough to represent most appraisal theories. Thus, those appraisal theories that argue for a relation event-appraisal variables [Ortony *et al.*, 1988; Scherer *et al.*, 2001] fit in the appraisal process of the architecture, since it has both `external events` and `internal events` as inputs, as shown in Table 3.1. On the other hand, appraisal theories that use processes with several steps [Scherer, 2001; Smith & Lazarus, 1990], can also be represented in this architecture. In Richard S. Lazarus's theory [Smith & Lazarus, 1990], for example, *primary* and *secondary* appraisals are performed in

### 3.4. Agent Attributes for BDI and Affective Processes

---

two different stages<sup>1</sup>. During primary appraisal the relevance and congruence of an event is evaluated starting from the agents concerns, and current options. During secondary appraisal, the event is evaluated according to the agent’s capacities and resources.

One way of representing the agent’s capacities and resources could consists of using the probabilities of the agent’s beliefs in such a way that a belief  $b$  can represent how probable the agent thinks  $b$  is regarding its current capacities and resources. Other issues like the memory of affectively relevant events may influence the appraisal process as well [Damásio, 1994]. Our idea of memory is in line with the psychological concept of “autobiographic memory” in psychology [Nelson, 1993], and it stores meaningful experiences of an individual. Optionally, and depending on the structure of specialized versions of *GenIA*<sup>3</sup>, affectively relevant events can be updated. Thus, if some event causes a strong emotional impact, it can be saved as an event that has been significant from an affective point of view.

The main outputs of the appraisal process are the appraisal variables<sup>2</sup>. These variables vary from one computational appraisal model to another but they have in common that these variable represent the agent’s judgments, and they produce changes in the affective state. The way these changes are produced is determined by the *affect generator* process, which uses the appraisal variables and the current affective state as inputs and obtains a modified affective state, as shown in Table 3.1. The *coping* process acts as an affect regulator that generates new intentions or modify the agent’s beliefs in order to take the current affective state to a desired state or to an “equilibrium state”. The intentions generated in this process can be oriented to perform, for example, reactive physical actions such as body gestures or facial expressions as a way to mitigate negative emotions. Each agent may have

---

<sup>1</sup>Lazarus’s theory also argues that primary and secondary appraisals are not necessarily sequential, but we limit it to be sequential in our architecture in order to have as much compatibility with other theories as possible.

<sup>2</sup>Although appraisal variables are an output of the appraisal process, they are not considered in Table 3.1 because this table only represents inputs and outputs that are persistent during the agent’s life cycle or that are used by two or more processes in the architecture.

### 3. GenIA<sup>3</sup>

---

particular ways of reacting to changes on its affective state, so these individual “coping strategies” are part of the agent’s personality, and they are used by the coping process. In order to determine the way the agent’s beliefs can be modified to cope with certain appraisals or changes on the affective state, the coping process can invoke the *affective modulator of beliefs* process. This process contains the mechanisms whereby certain affective state may influence beliefs. Finally the *affect’s temporal dynamic* process, is in charge of controlling the duration and decay of the affective state. As shown in Table 3.1, it can also use the traits of the agent’s personality as an input (see, for example, computational approaches like [Santos *et al.*, 2011]). Certainly there is a need of a balance between the temporal dynamic and the structural dynamic of the affective state in GenIA<sup>3</sup>, so each specialized version of GenIA<sup>3</sup> should pay attention and take care of this issue. Similarly, specialized architectures need to care about the balance between the frequency of the affective and reasoning cycles, since this design allows to give different priorities to each cycle.

#### 3.4.3 Representation of Affect-Related Attributes in an Agent Language

After identifying the main processes for modeling affective agents, it is necessary to address what should be represented when programming the agent in order to make all this “machinery” work. Current BDI-based agent languages include well established structures to represent and manage beliefs, options (or desires), intentions, as well as states, actions, and internal and external events. Therefore we focus on the representation of the knowledge used by affect-related processes. In this direction the distinction is made in three stages: affect generation, experience and effect.

##### 3.4.3.1 Affect Generation

In our proposal affect generation is performed by the appraisal theory. In order to be able to perform the *appraisal* process in a way that is consistent with most appraisal theories, a BDI agent language needs new structures to represent some affect-related attributes. We summarized these attributes in

### 3.4. Agent Attributes for BDI and Affective Processes

---

concerns, probabilities of beliefs, and personality following the conceptual system presented in Section 2.4.1.

**Concerns.** One of the most recurrent ideas on appraisal theories is the important role of individual concerns (e.g., motivations, standards, ideals) in the agent’s appraisal of events. According to Nico H. Frijda “a concern is what gives a particular event its emotional meaning” [Frijda, 2007]. Appraisal variables like “desirability of an event” of Andrew Ortony’s *et al.* appraisal theory [Ortony *et al.*, 1988], or “motivational relevance” and “motivational congruence” of Craig A. Smith’s and Lazarus’s theory [Smith & Lazarus, 1990], depend on the agent’s concerns.

**Personality.** In most affect-related theories, as well as the computational approaches that are based on them, the need of individual characteristics that differentiate the behavior of agents is addressed. Thus we believe that this is an important affective concept that needs a representation in an agent language. Personality can take several forms. Individual traits can be represented through dimensions (in line with dimensional theories of personality [Eysenck, 1990; McCrae & John, 1992; Myers *et al.*, 1985]). Also, personality may include concepts like coping strategies or rationality level. Coping strategies define the way an individual reacts to an event that involves emotional changes, and those reactions can be either involuntary manifestations or more planned actions. Reactions can be oriented to change beliefs, or intentions of an individual. We also propose the rationality level as part of the agent personality. It allows to establish mechanisms to implement agents that can be more or less emotional when making decisions according to their personality.

**Belief probability.** Computational approaches that model affective agents, often use formalization theories based on BDI logics, since “most (at least most ‘higher’) emotions are thought to depend on beliefs and—directly or indirectly—on desires” [Reisenzein *et al.*, 2013]. Nevertheless emotions also arise as the result of constructed representations of “what it could

be” either in the present or in the future (also defined as *prospects based* emotions by Ortony *et al.* [1988]). This is the case of, for example, *hope* or *fear*. In [Steunebrink *et al.*, 2012] *hope* is formalized as “being pleased about a prospective consequence (of an event)” and *fear* is formalized as “being displeased about a prospective consequence (of an event)”. Also by being able to compare reality with “what it could have been, might have been, or should have been” [Kahneman & Miller, 1986], an agent may experience emotions like *frustration* or *regret*. The representation of these future states (such as “what it could be”) can be done by assigning probabilities to beliefs for a given time  $t$  in the future. This representation also allows to evaluate the probability of the state of the world (that those beliefs conform) at time  $t$ . Thus, the agent language that supports approaches using these kind of emotions, should have useful structures for agents to perform “prospective reasoning” and for representing future states. An agent able to perform “prospective reasoning” can evaluate what a future affective state could be.

#### 3.4.3.2 Affect Experience

In order to associate the agent’s actions to specific affective states, the *affective state* needs a representation in the agent language. This representation can adopt a qualitative or quantitative way of expressing values of individual appraisal variables, emotion categories, values for dimensions of mood, or a combination of them.

#### 3.4.3.3 Affect Effects

In *GenIA*<sup>3</sup> the influence of affect on the agent’s behavior can be of two kinds: (1) by leaning the agent’s decisions towards options biased from what can be considered as ‘best rational options’, and (2) through responses of the agent in order to either adequate its emotional state for it to be more ‘desirable’, or to make the current situation become more controllable. The functions that are needed for the first kind of influence are performed by the *filter* process of

*GenIA*<sup>3</sup>, and the functions for the second kind are performed by the *coping* process. Ortony called these responses: “response tendencies” [Ortony, 2003]. He grouped the response tendencies in behavioral (automatic and difficult to control), information-processing (involuntary responses that produce changes on the way information is processed), and coping (strategies that transform the emotional state or the situation to be more favorable). Each particular agent language that complies with *GenIA*<sup>3</sup> should provide structures to allow the affective state to generate action tendencies by assigning “states of readiness” to actions or plans [Frijda, 1987], and should provide structure to define specific coping strategies.

## 3.5 Conclusions

An architecture like *GenIA*<sup>3</sup> offers a feasible and comprehensive way of developing affective agents by unifying different emotion theories on the basis of a common conceptual system. It facilitates the creation of computational models that focus on specific phenomena of interest, by relieving the modeler of irrelevant implementation choices or design specifications, and providing plausible default values. Hence, by proposing *GenIA*<sup>3</sup>, the Goal G2.1 of this thesis has been achieved.

Although similar architectures have been proposed (e.g., the architectures FAtiMA Modular [Aylett & Paiva, 2012; Dias *et al.*, 2014] and MAMID [Hudlicka, 2004, 2008] described in Chapter 2), *GenIA*<sup>3</sup> has been conceived with a different perspective. Instead of a modular approach, it is part of a toolbox of basic theory-elements that can be extended and/or replaced in order to fulfill specific requirements, allowing to model different affective phenomena with different psychological theories. Besides, in *GenIA*<sup>3</sup>, due to its inner structure, it is possible to set an equilibrium between the rational and the affective sides, which is the Goal G2.5 of this thesis, and which has been highlighted in the literature as a relevant issue of human behavior<sup>1</sup> [De Sousa, 1990]. *GenIA*<sup>3</sup> not only offers mechanisms for specifying how the affective

---

<sup>1</sup>Specifically the prevalence, on each individual, of either emotions or rationality on his behavior, or the relation between these two aspects.

### 3. GenIA<sup>3</sup>

---

state may influence behavior, but also mechanisms for specifying how changes on the affective state generate new behaviors through an explicit and integrated way of defining a coping model.

This chapter presented *GenIA*<sup>3</sup> at an abstract level, without making any commitment regarding to the theories that each one of its parts supports. In Chapter 4 the details of a default design of *GenIA*<sup>3</sup> is presented, as well as the formalization of its affective agent language.



---

## Formalization of GenIA<sup>3</sup>

### 4.1 Introduction

In Chapter 3, the main components of the *GenIA*<sup>3</sup> architecture were described. The present chapter takes up the Goals G2.2 to G2.5, and Goal G3 of this thesis presented in Chapter 1, by offering a more detailed description of *GenIA*<sup>3</sup>, and by making some commitments in a default design.

Specifically, this chapter proposes a formalization of *GenIA*<sup>3</sup> [Alfonso *et al.*, 2017] by extending the BDI agents' core structure (Goal G2.2). The reasoning cycle of a BDI agent is extended, and the agent is endowed with an affective cycle. Besides, the transition rules between the steps of the reasoning and affective cycles are proposed, by extending the operational semantics used in the Jason BDI agent platform [Bordini & Hübner, 2010; Vieira *et al.*, 2007]. The syntax of the Jason agent language is also extended, to include the agent's main attributes proposed in Chapter 3 (Goal G2.3). This chapter also describes the default design of *GenIA*<sup>3</sup> (Goal G2.4). This default design can be adapted in order to fulfill specific functionalities required by any application domain. Moreover, mechanisms for establishing an equilibrium between the rational and the affective sides of a *GenIA*<sup>3</sup> agent are also offered (Goal G2.5). These mechanisms allow adjusting the level of rationality of an individual, and the

frequency of rational and affective processes. These basic theory-elements have been implemented in the extension of the Jason's MAS platform that is also described in this chapter (Goal G3).

### 4.2 Extension of the Jason Agent Reasoning Cycle

Two alternatives can be explored in order to build human-like agents whose execution and representation consider both affective and rational processes. The first alternative is to build an agent engine and agent platform from scratch, and the second alternative is to extend an existing one. Considering that there exist well established and extensible agent engines and platforms that are compliant with the practical reasoning philosophy, we have chosen the second alternative. We have extended the reasoning cycle of a Jason agent [Bordini & Hübner, 2010; Vieira *et al.*, 2007] in order to build human-like agents whose execution and representation consider both affective and rational processes. Jason is a well-know agent-oriented programming language and platform, which is grounded in a logical computable language: AgentSpeak [Rao, 1996]. Jason is widely accepted in the agents' community due to its versatility to be adapted to several kinds of agent applications, thus, it becomes a suitable choice for building "customizable" affective agents able to represent a wide set of situations.

Hereinafter, we refer to *affective state* as a generalized representation of all agent attributes that characterize one or more aspects of the agent state in line with the definition of *core affect* of Russell [2003]: "A neurophysiological state that is consciously accessible as a simple, nonreflective feeling (...);" *affective processes* as the new processes added to the original BDI processes in order to consider affective characteristics; *affective cycle* as the cycle which modifies or generates the *affective state*; and *reasoning cycle* as the cycle that represents the agent practical reasoning. Moreover, we refer to the ranges of values for the variables that define the *affective state* as *affective categories*, and to a single emotion as an *emotion category* (e.g., joy, fear, or anger, in line with classifications like the one of the Ortony, Clore, and Collins' model (OCC) [Ortony *et al.*, 1988]).

## 4. Extension of the Jason Agent Reasoning Cycle

---

Figure 4.1 shows the steps of the three cycles that are part of the agent execution, as well as the relationship between these steps. The colored steps are either new or modified, while the non-colored steps maintain the same structure than those originally proposed in [Vieira *et al.*, 2007] with the same name. Similarly, the transitions with dashed lines are new or modified, and their corresponding transitions rules are presented in Section 4.4. There is a clear correspondence between these steps and the processes of *GenIA*<sup>3</sup>. The affective processes **appraisal**, **affect generator**, **coping**, **affective modulator of beliefs**, and **affect temporal dynamics** of *GenIA*<sup>3</sup> (see Section 3.2.2) are performed in the steps `Appr`, `UpAs`, `Cope`, `AffModB`, and `AsDecay` of Figure 4.1, respectively. The process `AffModB` has been integrated into the reasoning cycle, since it is closely linked to the addition and deletion of beliefs that take place in the reasoning cycle. The process `SelCs` is part of the **coping** process of the *GenIA*<sup>3</sup> architecture, and it is in charge of determining the coping strategies that need to be executed in the current affective cycle. The **brf** process is performed through the `Perceive`<sup>1</sup> and `ProcMsg` steps; the steps `SelEv`, `RelPl`, and `ApplPl` perform the **options** process; the **filter** process is performed through the `SelAppl`, `AddIM`, and `SelInt` steps; and the `ExecInt` and `ClrInt` steps are in charge of the **execute** process.

Figure 4.1 also shows that three possible cycles can take place during the agent execution. These cycles control the affective processes (affective cycle), the rational processes (reasoning cycle), and the temporal dynamics of affect (affect temporal dynamics cycle). Next, we describe the steps for each cycle. Before starting with the affective cycle it is worth mentioning that, in an initialization stage, the *affective state* has an initial value. This value is also the agent “equilibrium state”, which is a neutral state where the agent doesn’t experience any significant emotion. The affective cycle starts with the `Appr` step, where the **appraisal** process is performed by using several parameters including the agent *concerns*, *personality*, and the *probabilities* associated with

---

<sup>1</sup>Although the formalization of AgentSpeak considers that an agent can perceive new information from the environment, to the best of our knowledge, there is no explicit step in the reasoning cycle for this task. We have decided to make this step explicit with the initial step `Perceive`.

#### 4. FORMALIZATION OF GENIA<sup>3</sup>

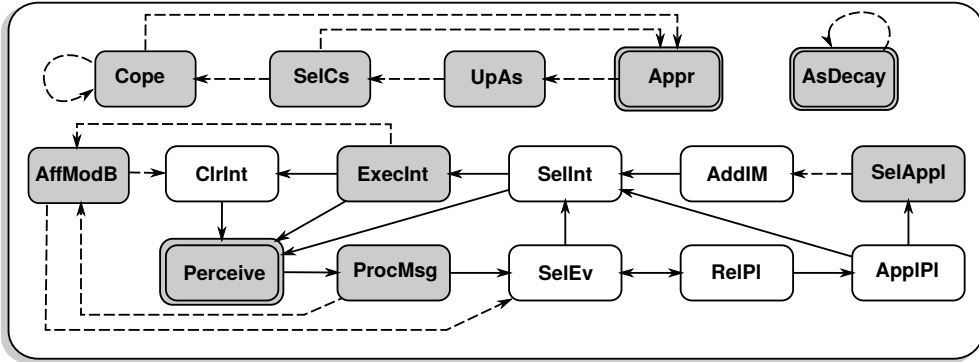


Figure 4.1: Extension of the reasoning cycle of AgentSpeak.

agent *beliefs* (e.g., if prospect-based emotions are generated). The `Appr` step also determines whether the appraised event is relevant for the agent from an affective point of view (in case the new *affective state* has an event associated). Then, in the `UpAs` step, the *affective state* is updated by using the appraisal variables generated in the `Appr` step. After the updating of the *affective state*, the `SelCs` step verifies whether it is necessary to generate new behaviors in the agent according to this change on the *affective state*, and verifies which of the agent *coping strategies* are applicable. The `Cope` step performs the tasks required to execute the selected coping strategies. The intentions derived from the execution of coping strategies are added as intended means at the end of the base of intentions, which is shared by both reasoning and affective cycles. Both reasoning and affective cycles generate their own intentions independently, which are included in this common base of intentions. Intentions generated by the affective cycle are added at the end of the current intentions, as well as the intentions generated by steps of the reasoning cycle. In the default design, all intentions are executed by their insertion order in the `SelInt` step of the reasoning cycle. The default implementation of this step is explained in Section 4.5.

The reasoning cycle contains two new steps, `Perceive` and `AffModB`, as well as three modified steps: `ProcMsg`, `SelAppl`, and `ExecInt`. `Perceive` is the initial step of the reasoning cycle. In this step, the agent *beliefs* are modified according to what can be observed from the environment as the result

of external events. The `Perceive` step is followed by the `ProcMsg` step, which is in charge of processing the messages received from other agents. Next, the information coming from the received messages, and from the perception of the environment (on the `Perceive` step), can be modified in the step `AffModB` (according to the agent's *affective state*), which follows the `ProcMsg` step. In *GenIA*<sup>3</sup> the agent's *affective state* also influences the agent's decisions. The `SelApp1` step performs this task by selecting the next applicable plan; thus, it has been modified to consider the agent current *affective state*, and also the agent *rationality level*. The step `ExecInt` may require to execute intentions that imply adding or removing beliefs. Thus, it can also be followed by the step `AffModB` to this end. Finally, the affect temporal dynamics cycle contains one step: `AsDecay`. This step determines the tendency of the *affective state* to return to its "equilibrium state". This task could use some traits of the agent *personality*.

### 4.3 Extension of the Jason Language

In Section 4.2, new cycles containing new steps were introduced in the `AgentSpeak`'s agent execution, and some existing steps were modified as well. These steps use some parameters that are agent attributes, which do not have a corresponding representation in the `AgentSpeak` agent language. For example, aspects of the agent *personality*, the agent *affective state*, the agent *concerns*, and the *probabilities* of the beliefs<sup>1</sup> are used in one or more steps. In this section a formal syntax for these attributes, is presented as an extension of the Jason agent language. In order to extend the specification of the Jason agent language, we extend the syntax in Extended Backus–Naur Form (EBNF) used in [Bordini *et al.*, 2007] (see Figures 4.2 and 4.3). Figure 4.2 shows an extract of the new syntax of the agent program, and Figure 4.3 shows an extract of the new syntax of the Multi-Agent System (MAS) project configuration. The next illustrative example will be used along this section to describe each new element of our extended version of the Jason language:

---

<sup>1</sup>In the default design of *GenIA*<sup>3</sup> these probabilities are interpreted as the agent's expectations (see Section 4.5).

#### 4. FORMALIZATION OF GENIA<sup>3</sup>

---

<u>agent</u>	→ <u>init_bels</u> [ <u>concerns</u> ] [ <u>personality</u> ] <u>init_goals</u> <u>plans</u>
<u>init_bels</u>	→ <u>beliefs</u> <u>rules</u>
<u>beliefs</u>	→ ( ( <u>literal</u>   <u>literal_prob</u> ) “.” )*
<u>concerns</u>	→ “concerns__” “(” <VAR> “)” “:-” <u>logical_expr</u> “.”
<u>personality</u>	→ “personality__” “:” “{”traits [“,” <u>rat_level</u> ] [“,” <u>coping_strats</u> ]“}” “.”
<u>traits</u>	→ “[”<NUMBER>“(“,”<NUMBER>)*”]”
<u>rat_level</u>	→ <NUMBER>
<u>coping_strats</u>	→ “copingst__” “:” (cs)+
<u>cs</u>	→ “cs_” “(” <u>logical_expr</u> “,” <u>aff_categ</u> “)” “-” <u>body</u> “.”
<u>init_goals</u>	→ ( “!” <u>literal</u> “.” )*
<u>plans</u>	→ ( <u>plan</u> )*
<u>plan</u>	→ [ “@” <u>atomic_formula</u> ] <u>triggering_event</u> [ “:” <u>context</u> ] [ “<-” <u>body</u> ] “.”
<u>literal_prob</u>	→ [ “~” ] <u>atomic_formula_prob</u>
<u>atomic_formula_prob</u>	→ ( <ATOM>   <VAR> ) [ “(” <u>list_of_terms</u> “)” ] [ “[” <u>list_of_terms</u> “]” ] [ <u>time_point_range</u> ]
<u>aff_categ</u>	→ <ATOM>
<u>time_point_range</u>	→ “<” <u>arithmetical_expr</u> “:” <u>arithmetical_expr</u> “>”

Figure 4.2: Simplified extension of the EBNF for the Jason Agent Language.

**Example 2** Consider an *extroverted* and *conscious* student who wants to pass an English exam. He is also interested in getting the maximum score in Calculus and he is pretty sure that he will get it. Scores range from 0 to 5, and the minimum required score to pass an exam is 3. Besides, normally the student’s affective state almost doesn’t influence his decisions.

Example 2 brings up two characteristics of the student, his concerns and his personality. The syntax of the Jason language has been extended in order to include these two new components, in addition to the agent’s initial beliefs (init\_bels), initial goals (init\_goals), and plans, as shown in the rule for agent in Figure 4.2. Thus, when writing the agent code, the programmer is now able to define the agent’s concerns<sup>1</sup> and personality.

---

<sup>1</sup>This view of concerns can be confused with declarative goals. Unlike declarative goals, which are goals that explicitly represent a state of affairs to be achieved, concerns are a way

In Example 2, concerns and personality can be defined as follows:

```
concern1(C1):- score(SEng,english) & ((SEng >= 3 & C1 = 1) |
    (SEng < 3 & C1 = 0)).
concern2(C2):- score(SCal,calculus) & ((SCal == 5 & C1 = 1) |
    (SCal < 5 & C1 = 0)).
concerns__(V):- concern1(C1) & concern2(C2) & V=(C1+C2)/2.
personality__ : { [0.3, 0.9, 0.85, 0.6, 0.1], 0.7 }.
```

Observe that the value of concerns range from 0 to 1, and that this value is stored in the `V` variable. In this example, the agent's personality is compliant with the Five Factor Model of personality [McCrae & John, 1992], whose components are 'openness', 'conscientiousness', 'extraversion', 'agreeableness', and 'neuroticism' in this order. As the student is highly conscious and extroverted, the second and third values of the list of values of the personality traits are high values in a range from 0 to 1 (0.9 and 0.85 respectively). Besides, as the student is pretty rational, because his affective state almost doesn't influence his decisions, he has a fairly high level of rationality (with a value of 0.7).

We also propose new structures in order to express the probabilities of the agent's beliefs. As shown in Figure 4.2, in the new belief syntax, each belief can be written either in the traditional way or with an annotation indicating its probability. We defined the new reserved word `prob__` that represents the functor of a term (with one numerical term as argument), which indicates the probability of the corresponding belief (when the term is one of the annotations of the belief). This new kind of belief can also has a new component, which indicates the temporal interval that the agent thinks this belief holds and/or will hold. If a belief doesn't have any probability annotation, the default probability 1 will be used; if a belief doesn't have a temporal interval, it expresses that the belief holds in that moment. In the reasoning process, probabilities are propagated following the Bayes theory. The start and end time points of the temporal interval can be defined as a number, a reserve

---

of evaluating how "good" the state of affairs is. In other words, they offer a direct assessment of the state of affairs without needing to represent declarative goals. For example, a situation where an agent has declarative goals, and where the current state of affairs doesn't satisfy any of them, would be represented with a very low value of the agent's concerns.

## 4. FORMALIZATION OF GENIA<sup>3</sup>

---

word (`Now` or `Infinite`), or an arithmetical expression that can contain the reserved word `Now`. In Example 2, the student’s expectation about the maximum score in Calculus can be expressed as one of the following ways:

```
score(5, calculus) [prob__(0.95)] <Now, Infinite>
score(5, calculus) [prob__(0.95)] <Now, 24*60*60+Now>
```

The first way expresses the student’s expectation of getting the maximum score in all future exams of Calculus with a probability of 0.95. The second way, in turn, expresses the expectation of getting the maximum score in Calculus in the next 24 hours with the same probability.

We also defined the reserved word `affect__` as the functor of a literal with one or more “Atom” terms as argument. This arguments represent affective categories that must match with one of the `affect_categories` included in the project configuration (see Figure 4.3). This literal can be used as a plan annotation, and it will be considered as an additional criteria in the selection of an applicable plan. Suppose that the affective state of the student of Example 2 has two components: pleasure and arousal. Suppose also that there exists an affective category ‘excited’ that becomes active for specific values of pleasure and arousal. The structure of a plan for celebrating after knowing a good score can be expressed as follows:

```
@p_celebrate[affect__(excited)]
+score(5, Subject)
<- .print("I celebrate the maximum score in ", Subject);
    (...).
```

The previous extract of code shows a plan `p_celebrate` that will have priority over other plans with the same triggering event, when the agent’s affective state matches the affective category `excited`.

The syntactic specification for the MAS project configuration in Figure 4.3 has been extended in two directions. First, the agent specification includes three new optional components: `emEngineClass`, `personalityClass`, and `affStateClass`. The reserved word `emEngineClass` allows to assign the



## 4. Extension of the Jason Language

<u>mas</u>	→ “MAS” <ID> “{” [ <u>infrastructure</u> ] [ <u>environment</u> ] [ <u>exec_control</u> ] [ <u>affect_categories</u> ] “}”
<u>agent</u>	→ <ASID> [ <u>filename</u> ] [ <u>options</u> ] [“beliefBaseClass” <ID>] [“agentClass”<ID>] [“agentArchClass” <ID>] [“emEngineClass” <ID>] [“personalityClass” <ID>] [“affStateClass” <ID>] [“#” <NUMBER>] [“at” <ID>] “;”
<u>affect_categories</u>	→ “affect_categories” “:” <u>aff_category</u> ( “,” <u>aff_category</u> )* “;”
<u>aff_category</u>	→ <ATOM>“ (“ <u>aff_components</u> “)”
<u>aff_components</u>	→ ( (<NUMBER> “-”<NUMBER>)   <u>range</u> ) ( “,” (<NUMBER> “-”<NUMBER>)   <u>range</u> )*
<u>range</u>	→ “<” (<NUMBER> “-”<NUMBER>) “:” (<NUMBER> “-”<NUMBER>) “>”

Figure 4.3: Simplified extension of the EBNF of the MAS project configuration.

emotion engine class to the agent, which contains the functions that implement the main affective processes; `personalityClass` allows to assign to the agent the class that manages its personality; and `affStateClass` allows to assign to the agent the class that manages its affective state. Besides, it is possible to use a new reserved word in the agent options, which is `nacprc` (number of affective cycles per reasoning cycle). This option allows to establish the relation between the affective and reasoning cycles (by default it is 1:1). Second, the new component `affect_categories`, represents a list of affective categories. Next, an extract of a project configuration with a single agent representing the student of Example 2 is shown.

### agents:

```
student [nacprc=10:3]
  agentClass jason.Em.asSemantics.EmAgent
  agentArchClass jason.Em.architecture.EmAgArch
  emEngineClass jason.Em.architecture.DefaultEmotionEngine
```

## 4. FORMALIZATION OF GENIA<sup>3</sup>

---

```
personalityClass jason.Em.architecture.OCEAN
affStateClass jason.Em.architecture.PAD;
affcategories:
  excited(<0.1:1.0>,<0.8:1.0>),
  bored (<0.0:0.4>,<0.0:0.2>),
  happy (<0.8:1.0>,<0.6:1.0>),
  sad (<0.0:0.2>,<0.0:0.2>);
```

This extract of code shows the configuration of a MAS with a single ‘student’ agent whose execution has 10 affective cycles each 3 reasoning cycles. Also, the agent’s definition, the architecture configuration, as well as the emotion engine, personality, and affective state definitions are performed through the classes EmAgent, EmAgArch, DefaultEmotionEngine, OCEAN, and PAD respectively. Besides, there have been defined four affective categories in the project (exited, bored, happy, and sad), that can be used for the plans’ prioritizing. Each of these affective categories has two ranges of values for the components of the affective state pleasure and arousal.

Table 4.1 shows a more detailed explanation of some of the components of Figures 4.2 and 4.3. The grammars presented in this section have been simplified and adapted for them to be more intelligible. Sections B.2 and B.3 show the full version of the grammars used to generate the parsers used by our Jason extension, and Sections B.4 and B.5, show examples using these grammars.

### 4.4 Extension of the AgentSpeak’s Operational Semantics

In Section 4.3, an extension of the Jason language with new affect-related attributes was presented. This section presents the transitions rules that describe the use of these attributes on the steps of the Jason agent reasoning and affective cycles, as well as the transitions between these steps. This transition rules are part of the Jason operational semantics.

In order to build an extension of the Jason platform, whose agents try to simulate a human-like behavior, with affect-related processes and characteristics,

## 4. Extension of the AgentSpeak’s Operational Semantics

Table 4.1: Explanation of the main components of the EBNF extension for a Jason agent.

	Concept	Involved components and meaning
Agent EBNF	concerns	<p><b>concerns</b>: A rule, whose left side is a literal, which contains one numerical term, indicating the extent to which the current beliefs are aligned with the agent concerns. The right side of the rule expresses the way this numerical value is calculated according to the agent’s current beliefs.</p>
	personality	<p><b>personality</b>: The agent personality. It has three components: <b>traits</b>, <b>rat_level</b>, and <b>coping_strats</b>. The last two are optional.</p> <p><b>traits</b>: In line with the traits perspective of personality [Allport, 1937; Eysenck, 1982; Goldberg <i>et al.</i>, 1990]. It is a list of numerical values, each representing the extent to which its corresponding <b>personality</b> trait is present in the agent personality.</p> <p><b>rat_level</b>: A numerical value that indicates the extent to which an individual is more rational and, hence, less emotive when making decisions.</p> <p><b>coping_strats</b>: Represents the agent set of coping strategies.</p> <p><b>cs</b>: A single coping strategy. Its left side has two components: a context and an affective category (<b>aff_categ</b>), and its right side represents a list of actions<sup>a</sup> (which represent the agent behavior for this coping strategy). Actions in <b>cs</b>, are executed if the <b>context</b> is a logical consequence of the agent belief base and if the agent current affective state matches <b>aff_categ</b><sup>b</sup>.</p> <p><b>aff_categ</b>: A label indicating one of the affective categories defined in the project configuration (<b>affect_categories</b>).</p>
	beliefs probabilities	<p><b>literal_prob</b>: A literal for representing beliefs with probabilities. Its atomic formula (<b>atomic_formula_prob</b>) contains a time range.</p> <p><b>time_point_range</b>: Temporal interval that the agent thinks the corresponding belief holds and/or will hold.</p> <p><b>time_point</b>: Represents the lower or upper bound of a time range.</p>
Project Configuration EBNF	affect	<p><b>affect_categories</b>: Represents a set of affective categories.</p> <p><b>aff_category</b>: A single affective category. It has a functor and a list of components, where each component can be either a numerical value or a range of values. These ranges contain the values that the affective state should have to match this affective category<sup>c</sup>.</p> <p><b>range</b>: A range that contains the minimum and maximum values for each affective component.</p>

<sup>a</sup> The right side of a coping strategy is similar to the body of a Jason plan (as shown in the extension of the EBNF of Jason in Figure 4.2). The body of a Jason plan can include conditional statements, loops, and so on.

<sup>b</sup> Due to their nature, coping strategies are not processed as Jason plans or rules. That’s why a special structure has been created for them.

<sup>c</sup> The number of components of an affective category (and, hence, of an affective state), depends on the psychological theory used.

#### 4. FORMALIZATION OF GENIA<sup>3</sup>

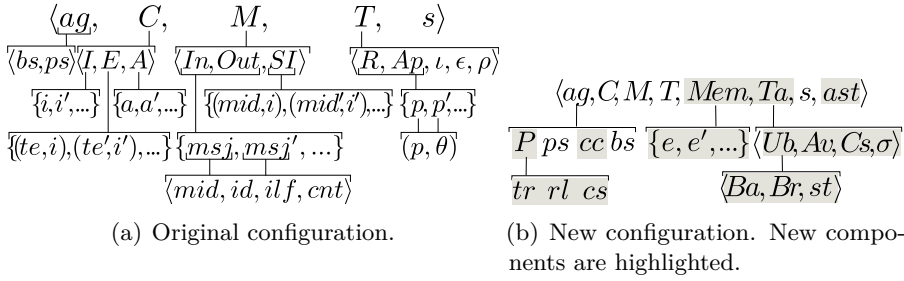


Figure 4.4: Configuration of a Jason agent.

we extended the AgentSpeak’s operational semantics, considering that it is the base of the Jason’s operational semantics. This operational semantics describe the changes of the agent’s configuration on each transition. In AgentSpeak this agent configuration is defined as a tuple  $\langle ag, C, M, T, s \rangle$ . Figure 4.4(a) shows this configuration as well as the structure of each one of its components. The new Jason agent configuration has the form  $\langle ag, C, M, T, Ta, Mem, s, ast \rangle$ . Figure 4.4(b) highlights the new components of this configuration, which are described next.

- $ag$  represents the agent program, which originally contains a set of beliefs ( $bs$ ), and a set of plans ( $ps$ ). Additionally, a set of concerns ( $cc$ ), and a personality ( $P$ ) has been included in the agent program  $ag$ .
  - The agent concerns  $cc$  is an agent attribute which is in line with the concerns in *GenIA*<sup>3</sup>, and reflects the agent’s ideals, motivations, needs, interests, and/or standards.
  - Personality  $P$  includes  $tr$ ,  $lr$ , and  $cs$ .  $tr$  contains a set of numerical values representing the agent personality traits. For instance, the Five Factor Model of personality [McCrae & John, 1992] argues that the traits openness, conscientiousness, extraversion, agreeableness, and neuroticism can differentiate an individual from the rest. In line with the “rationality level” of *GenIA*<sup>3</sup> we propose  $rl$ , which is also part of the agent personality. The rationality level states the extent to which the agent’s decisions are influenced by its affective state.

## 4. Extension of the AgentSpeak's Operational Semantics

---

*cs* represents the agent coping strategies, which relate a particular state (represented through a set of beliefs) and an affective state with a set of actions that generate intentions to be included in the agent's current intentions.

- *C*, *M*, and *T* were originally part of AgentSpeak and represent the agent circumstance, communication parameters, and temporary information for a reasoning cycle.
- *Mem* contains a set of events  $\{e, e', \dots\}$  that have been relevant for the agent from an affective point of view. We consider this set as a kind of “autobiographic memory”, where the meaningful experiences are stored, as proposed by Nelson [1993]. These events are determined and *Mem* is updated in the appraisal process.
- *Ta* is a tuple  $\langle Ub, Av, Cs, \sigma \rangle$ , which represents the temporary information used by the affective processes in an affective cycle. Its components are:
  - *Ub* is a tuple  $\langle Ba, Br, st \rangle$  which contains those beliefs to be added to or removed from the agent's belief base. *Ba* and *Br* represent the set of beliefs to be added and the set of beliefs to be removed, respectively; *st* contains the label of the step that requires to add and/or to remove beliefs in *Ba* and *Br*<sup>1</sup>.
  - *Av* contains the set of numerical values for the appraisal variables in the current affective cycle.
  - *Cs* contains the set of coping strategies to be executed in the current affective cycle.
  - $\sigma$  represents the agent's current affective state. It contains a set of variables  $\{v, v', \dots\}$  where each variable contains a numerical value representing the intensity or the presence or not (in the case a bivalent variable) of either an emotion category (e.g., sad, happy, angry), an appraisal variable (e.g., desirability, controllability), or a

---

<sup>1</sup>*Ub* is used by the step *AffModB*.

## 4. FORMALIZATION OF GENIA<sup>3</sup>

---

mood dimension (e.g., the dimensions of the PAD model [Mehrabian, 1996b]).

- $s$  is a label annotating the current step in the reasoning cycle, where  $s \in \{Perceive, ProcMsg, SelEv, RelPl, ApplPl, SelAppl, AddIM, SelInt, ExecInt, ClrInt, AffModB\}$ . A new label  $AffModB$  has been included in  $s$ , which corresponds to the new step  $AffModB$  in the reasoning cycle (see Figure 4.1).
- $ast$  is a label annotating the current step in the affective cycle, where  $ast \in \{Appr, UpAs, SelCs, Cope\}$  (see Figure 4.1).

Using a similar notation to that used in [Vieira *et al.*, 2007], we refer to attributes with a subindex. For example we refer to the appraisal variables  $Av$  that are part of the affective temporary information  $Ta$ , as  $Ta_{Av}$ . Similarly, we refer to the traits  $tr$  of the agent personality  $agp$  as  $agp_{tr}$ . We have also defined the structure of new functions that are part of the agent configuration and whose content must be specified by the agent programmer<sup>1</sup>. By offering a way of customizing these functions, the possibility of implementing various psychological theories is offered, thus fulfilling the Goal G2.1 of this thesis. An example of how these functions can be implemented is described in more detail in Section 4.5. One of these functions is  $AsDec(\sigma, P_{tr})$ , which controls how the affective state  $\sigma$  decays over time<sup>2</sup>, and obtains new values for this affective state considering the personality traits  $P_{tr}$ . Appraisal variables are generated through the function  $Appraise(\varepsilon, bs, cc, Mem, Ap)$ , which considers the event to evaluate ( $\varepsilon$ )<sup>3</sup>, a set of beliefs ( $bs$ ), concerns ( $cc$ ), the affective relevant events for the agent ( $Mem$ ), and options that the agent has available ( $Ap$ ). The function  $AffRelEv(\varepsilon, Av)$  evaluates if the event  $\varepsilon$  is relevant for the agent

---

<sup>1</sup>The selection of the type and number of the parameters of these functions is based on those most commonly used in related computational approaches, and not all of them should be necessarily used in every case.

<sup>2</sup>By default it is considered that the affective state decays, because in general psychological theories argue for this decay when dealing with the affect temporal dynamic [Frijda, 2007]. Nevertheless, the function  $AsDec(\sigma, P_{tr})$  can be customized to include any other behavior.

<sup>3</sup>Events in AgentSpeak include the addition and deletion of beliefs (from the environment or own), addition of goals, and failure of goals.

## 4. Extension of the AgentSpeak's Operational Semantics

---

from an affective point of view by using the appraisal variables in  $Av$ . The function  $\text{UpAffSt}(\sigma, Av)$  determines a new affective state (which contains a set of variables), given the affective state  $\sigma$  and according to a set of appraisal variables  $Av$ . Function  $\text{modB}(AddB, DelB, \sigma)$  determines what beliefs from the set  $AddB$  need to be added, what beliefs from the set  $DelB$  need to be removed, and what beliefs of both sets need a modification, according to the affective state  $\sigma$ .

We have also modified the AgentSpeak selection function  $S_{Ap}$  for selecting an applicable plan from the set of applicable plans  $Ap$ . It has the form  $S_{Ap}(Ap, \sigma, P_{rl})$ , where the affective state  $\sigma$  and an agent rationality level  $P_{rl}$  are new parameters. Besides, a new selection function  $S_{cs}(Cs)$  has been created in order to select a coping strategy from a set of coping strategies  $Cs$ . The selection functions  $S_{Ap}$  and  $S_{cs}$  are defined at design time by the agent programmer, according to the desired behaviors for the agents. We do not include the selection functions in the configuration for a better readability. Nevertheless, Section 4.5 offers an informal description of them in our default design. Finally, although not explicitly defined in all cases, the agent's personality traits  $P_{tr}$  can, in fact, be used in every function described in order to represent individual differences in the agent's processes. Due to this, personality traits are an agent attribute that is accessible by any process.

Additionally, we have defined the  $\text{EvalP}(PSet, bs)$ ,  $\text{match}(\sigma, ac)$ , and  $\text{SelCopeSt}(P_{cs}, bs, \sigma)$  functions to determine: changes on percepts of the environment, whether a particular affective state matches an affective category, and applicable coping strategies, respectively. Definitions 4.2, 4.3, and 4.4 propose a formalization for these functions.  $\text{agperc}(bs)$  of Definition 4.1 is an auxiliary function that determines the agent's current percepts.

**Definition 4.1** *Given the set  $bs$  of agent beliefs, the set of beliefs that correspond to the agent's percepts is defined as follows:*

$$\text{agperc}(bs) = \{b[\text{annot}] \mid b[\text{annot}] \in bs \text{ and } \text{source}(\text{percept}) \in \text{annot}\}$$

**Definition 4.2** *Let be  $bs$  the set of agent beliefs, and let be  $PSet = \{pc, pc', \dots\}$  the set of percepts observable in the environment (where each  $pc$  is a literal). The*

#### 4. FORMALIZATION OF GENIA<sup>3</sup>

---

set of new percepts  $NewP$  is calculated as the set difference  $PSet \setminus agperc(bs)$ . Also the set  $RemP$  of percepts no longer existing in  $bs$ , is calculated as the set difference  $agperc(bs) \setminus PSet$ . The function  $EvalP(PSet, bs)$  performs this task and it is defined as follows:

$$EvalP(PSet, bs) = \langle NewP, RemP \rangle \mid NewP = \{b \in PSet \mid b \notin agperc(bs)\} \\ \text{and } RemP = \{b \in agperc(bs) \mid b \notin PSet\}$$

**Definition 4.3** Let be  $\sigma = \{a_1, a_2, \dots, a_k\}$  be a set of  $k$  numerical values, each corresponding to an affective label, and let be  $ac = \{r_1, r_2, \dots, r_k\}$  be a set of  $k$  ranges of values for the same affective labels, where  $r_i = [rmin_i, rmax_i]$ . The  $match(\sigma, ac)$  function determines whether a particular affective state  $\sigma$  matches an affective category  $ac$ , and it is defined as follows:

$$match(\sigma, ac) = \begin{cases} \text{TRUE} & \text{if } a_i \geq rmin_i \text{ and } a_i \leq rmax_i \quad \forall i \in \mathbb{Z} \mid 1 \leq i \leq k \\ \text{FALSE} & \text{otherwise} \end{cases}$$

As shown in Section 4.3, a coping strategy has three components: *context*, *affective category*, and *body*. Both *context* and *body* have the same meaning and structure as the context and body of a plan, where *context* represents a set of conditions that must hold, and *body* contains a set of actions to be performed [Vieira *et al.*, 2007]. Also, to execute the *body*'s actions, the agent's current affective state must match the *affective category*. If a coping strategy  $cs$  has the form  $(ct, ac) \rightarrow h$ , where  $ct$  is the context,  $ac$  is the affective category, and  $h$  is the body, the function  $SelCopeSt(P_{cs}, bs, \sigma)$  is defined as follows:

**Definition 4.4** Let be  $P_{cs}$  a set of coping strategies,  $bs$  a set of beliefs, and  $\sigma$  a particular affective state. The set of applicable coping strategies is defined as follows:

$$SelCopeSt(P_{cs}, bs, \sigma) = \{(cs, \theta) \mid cs \in P_{cs} \text{ and } \theta \text{ is s.t. } bs \models ct\theta \\ \text{and } match(\sigma, ac) \\ \text{where } ct = CsCtxt(cs) \text{ and } ac = CsAc(cs)\}$$



## 4. Extension of the AgentSpeak's Operational Semantics

---

In the Definition 4.4, the functions  $CsCtxt(cs)$  and  $CsAc(cs)$  return the context and the affective category of a given coping strategy  $cs$ , and  $\theta$  is the most general unifier.

### 4.4.1 New Transition Rules

In this section, we present the transition rules for the updated or new steps of the agent cycles (see Figure 4.1), with respect to the steps on [Vieira *et al.*, 2007], using the Structural Operational Semantics (SOS) [Plotkin, 1981]. Note that the initial state of the reasoning cycle is  $\langle ag, C, M, T, Mem, Ta, Perceive, Appr \rangle$ . At this point, the steps of the cycles are able to update one or more components of the agent configuration. Next, we describe these transition rules individually. We start by the transition rules for the steps of the affective cycle.

**Appraisal** The process of appraisal takes place in these transition rules through the function  $Appraise(T_\varepsilon, ag_{bs}, ag_{cc}, Mem, T_{Ap})$ , which evaluates the current event  $T_\varepsilon$ . If the function  $AffRelEv(T_\varepsilon)$  returns TRUE, the current event  $T_\varepsilon$  is added to the set of affectively relevant events  $Mem$  (rule  $Appr_1$ ). The next step in this transition is the  $UpAs$  step.

$$\frac{AppVar = Appraise(T_\varepsilon, ag_{bs}, ag_{cc}, Mem, T_{Ap}) \quad AffRelEv(T_\varepsilon, AppVar)}{\langle ag, C, M, T, Mem, Ta, s, Appr \rangle \rightarrow \langle ag, C, M, T, Mem', Ta', s, UpAs \rangle} \quad (Appr_1)$$

$$\text{where: } Ta'_{Av} = AppVar \\ Mem' = Mem \cup T_\varepsilon$$

$$\frac{AppVar = Appraise(ag_{bs}, ag_{cc}, T_\varepsilon, Mem, T_{Ap}) \quad \neg AffRelEv(T_\varepsilon, AppVar)}{\langle ag, C, M, T, Mem, Ta, s, Appr \rangle \rightarrow \langle ag, C, M, T, Mem, Ta', s, UpAs \rangle} \quad (Appr_2)$$

$$\text{where: } Ta'_{Av} = AppVar$$

#### 4. FORMALIZATION OF GENIA<sup>3</sup>

---

**Update Affective State** In this transition, the agent affective state is updated through the function  $\text{UpAffSt}(Ta_\sigma, Ta_{Av})$ . The next step after this transition is the  $\text{SelCs}$  step.

$$\frac{\langle ag, C, M, T, Mem, Ta, s, \text{UpAs} \rangle}{\langle ag, C, M, T, Mem, Ta', P, s, \text{SelCs} \rangle} \quad (\text{UpAffState})$$

$$\text{where: } Ta'_\sigma = \text{UpAffSt}(Ta_\sigma, Ta_{Av})$$

**Select Coping Strategies** In this transition, the agent applicable coping strategies are determined through the function  $\text{SelCopeSt}(ag_{P_{cs}}, ag_{bs}, Ta_\sigma)$ .  $Ta_{Cs}$  is updated with the result of  $\text{SelCopeSt}(ag_{P_{cs}}, ag_{bs}, Ta_\sigma)$  and the cycle goes on with the step  $\text{Cope}$  (transition rule  $\text{SelCs}_2$ ). If no coping strategy is applicable, the cycle returns to the step  $\text{Appr}$  (transition rule  $\text{SelCs}_1$ ).

$$\frac{\text{SelCopeSt}(ag_{P_{cs}}, ag_{bs}, Ta_\sigma) = \{\}}{\langle ag, C, M, T, Mem, Ta, s, \text{SelCs} \rangle \rightarrow \langle ag, C, M, T, Mem, Ta, s, \text{Appr} \rangle} \quad (\text{SelCs}_1)$$

$$\frac{\text{SelCopeSt}(ag_{P_{cs}}, ag_{bs}, Ta_\sigma) \neq \{\}}{\langle ag, C, M, T, Mem, Ta, s, \text{SelCs} \rangle \rightarrow \langle ag, C, M, T, Mem, Ta', s, \text{Cope} \rangle} \quad (\text{SelCs}_2)$$

$$\text{where: } Ta_{Cs} = \text{SelCopeSt}(ag_{P_{cs}}, ag_{bs}, Ta_\sigma)$$

**Cope** In the step  $\text{Cope}$  the function  $S_{cs}(Ta_{Cs})$  selects a coping strategy from the current set  $Ta_{Cs}$  of applicable coping strategies. A plan  $p$  is created whose head is a TRUE value and whose actions (which are the body  $h$  of the plan  $p$ ) are those of the selected coping strategy. The plan  $p$  and the unifier  $\theta$  are added as an intention to the set of current intentions  $C_I$  and the selected coping strategy is removed from the set of applicable coping strategies  $Ta_{Cs}$ . The intention added can lead to the addition or dropping of beliefs, goals, and to a variety of actions (in general, all actions that Jason allows to perform in a plan body). This step is repeated until  $S_{cs}(Ta_{Cs})$  doesn't return any coping strategy, and then the cycle goes on with the step  $\text{Appr}$ .

## 4. Extension of the AgentSpeak's Operational Semantics

---

$$\frac{S_{cs}(Ta_{Cs}) = (cs, \theta) \quad cs = (ct, ac) \rightarrow h}{\langle ag, C, M, T, Mem, Ta, s, Cope \rangle \rightarrow \langle ag, C', M, T, Mem, Ta', s, Cope \rangle} \quad (\text{Cope}_1)$$

where:  $p = \text{true} \leftarrow h$

$$C'_I = C_I \cup \{[p\theta]\}$$

$$Ta'_{Cs} = Ta_{Cs} \setminus \{(cs, \theta)\}$$

$$\overline{\langle ag, C, M, T, Mem, Ta, s, Cope \rangle \rightarrow \langle ag, C, M, T, Mem, Ta, s, Appr \rangle} \quad (\text{Cope}_2)$$

**Perceive** This is the initial step of the reasoning cycle. The agent checks the environment for determining changes on percepts ( $PSet$ ) through the function  $\text{EvalP}(PSet, ag_{bs})$ .  $NewP$  contains new percepts to be included in the agent's belief base  $ag_{bs}$ , and  $RemP$  contains percepts to be removed from the agent belief base  $ag_{bs}$ . The next step in the cycle is  $\text{ProcMsg}$ , and both  $NewP$  and  $RemP$  are stored in the affective temporal information of the agent configuration as  $Ta_{Ub}$  for them to be processed later in the step  $\text{AffModB}$ .

$$\frac{\text{EvalP}(PSet, ag_{bs}) = \langle NewP, RemP \rangle}{\langle ag, C, M, T, Mem, Ta, Perceive, ast \rangle \rightarrow \langle ag, C, M, T, Mem, Ta', ProcMsg, ast \rangle} \quad (\text{Perc})$$

$$Ta'_{Ub} = \langle NewP, RemP, Perceive \rangle$$

The next four rules are related to the processing of received messages. In these rules, the functions  $S_M(M_{In})$  and  $\text{SoCAcc}(id, ilf, at)$ , and the set  $M_{SI}$  are used. The first function selects a message from the messages set  $M_{In}$ , and the second function determines if a message is “socially acceptable”, where  $id$  is the message identifier,  $ilf$  is the illocutionary force of the message, and  $at$  is the propositional content of the message. The set  $M_{SI}$  stores suspended intentions that are awaiting replies for previously sent messages. More details of these functions and set can be found in [Vieira *et al.*, 2007].

#### 4. FORMALIZATION OF GENIA<sup>3</sup>

---

**Receiving a Tell message** This transition has been modified in the same way as other transitions in which beliefs were added to the agent belief base. Thus, instead of adding them directly to the agent belief base, they are added to the affective temporal information of the agent configuration  $Ta_{Ub}$ , for them to be processed in the step  $AffModB$ .

$$\frac{S_M(M_{In}) = \langle mid, id, Tell, Bs \rangle \quad (mid, i) \notin M_{SI} \text{ (for any intention } i) \\ SocAcc(id, Tell, Bs) \quad Ta_{Ub} = \langle NewP, RemP, Perceive \rangle}{\langle ag, C, M, T, Mem, Ta, ProcMsg, ast \rangle \rightarrow \langle ag, C, M', T, Mem, Ta', AffModB, ast \rangle} \quad (\text{Tell}')$$

$$\text{where: } M'_{In} = M_{In} \setminus \{ \langle mid, id, Tell, Bs \rangle \}$$

$$Bs' = NewP$$

$$\text{and for each } b \in Bs: Bs' = Bs' \cup \{ b[source(id)] \}$$

$$Ta'_{Ub} = \langle Bs', RemP, ProcMsg \rangle$$

**Receiving a Tell message as Reply** Similarly, in this transition, beliefs sent by another agent as reply are added to  $Ta_{Ub}$  for them to be processed in the step  $AffModB$ . Also, the required actions to resume the required intention are performed.

$$\frac{S_M(M_{In}) = \langle mid, id, Tell, Bs \rangle \quad (mid, i) \in M_{SI} \text{ (for any intention } i) \\ SocAcc(id, Tell, Bs) \quad Ta_{Ub} = \langle NewP, RemP, Perceive \rangle}{\langle ag, C, M, T, Mem, Ta, ProcMsg, ast \rangle \rightarrow \langle ag, C', M', T, Mem, Ta', AffModB, ast \rangle} \quad (\text{TellRepl}')$$

## 4. Extension of the AgentSpeak's Operational Semantics

---

where:  $M'_{In} = M_{In} \setminus \{\langle mid, id, Tell, Bs \rangle\}$

$M'_{SI} = M_{SI} \setminus \{\langle mid, i \rangle\}$

$C'_I = C_I \cup \{i\}$

$Bs' = NewP$

and for each  $b \in Bs$ :  $Bs' = Bs' \cup \{b[\text{source}(id)]\}$

$Ta'_{Ub} = \langle Bs', RemP, ProcMsg \rangle$

**Receiving an Untell message** In this transition, beliefs that need to be removed as the result of a message of other agent are added to  $Ta_{Ub}$  for them to be processed in the step  $AffModB$ .

$$\frac{S_M(M_{In}) = \langle mid, id, Untell, ATs \rangle \quad (mid, i) \notin M_{SI} \text{ (for any intention } i) \quad \text{SocAcc}(id, Untell, ATs) \quad Ta_{Ub} = \langle NewP, RemP, Perceive \rangle}{\langle ag, C, M, T, Mem, Ta, ProcMsg, ast \rangle \rightarrow \langle ag, C, M', T, Mem, Ta', AffModB, ast \rangle} \quad (\text{Untell}')$$

where:  $M'_{In} = M_{In} \setminus \{\langle mid, id, Untell, ATs \rangle\}$

$DelB = \{at\theta \mid \theta \in \text{Test}(ag_{bs}, at) \wedge at \in ATs\}$

$DelB' = RemP$

and for each  $b \in DelB$ :  $DelB' = DelB' \cup \{b[\text{source}(id)]\}$

$Ta'_{Ub} = \langle NewP, DelB', ProcMsg \rangle$

**Receiving an Untell message as Reply** This rule is similar to the previous one, where beliefs that need to be removed as the result of a reply message of another agent are added to  $Ta_{Ub}$  for them to be processed in the step  $AffModB$ . Also, the required actions to resume the required intention are performed.

#### 4. FORMALIZATION OF GENIA<sup>3</sup>

---

$$\begin{array}{c}
 S_M(M_{In}) = \langle mid, id, Untell, ATs \rangle \quad (mid, i) \in M_{SI} \text{ (for any intention } i) \\
 \text{SoCAcc}(id, Untell, ATs) \quad Ta_{Ub} = \langle NewP, RemP, Perceive \rangle \\
 \hline
 \langle ag, C, M, T, Mem, Ta, ProcMsg, ast \rangle \rightarrow \langle ag, C', M', T, Mem, Ta', AffModB, ast \rangle \\
 \text{(UntellRepl')}
 \end{array}$$

$$\begin{array}{l}
 \text{where: } M'_{In} = M_{In} \setminus \{ \langle mid, id, Untell, ATs \rangle \} \\
 M'_{SI} = M_{SI} \setminus \{ (mid, i) \} \\
 C'_I = C_I \cup \{ i \} \\
 DelB = \{ at\theta \mid \theta \in \text{Test}(ag_{bs}, at) \wedge at \in ATs \} \\
 DelB' = RemP
 \end{array}$$

$$\begin{array}{l}
 \text{and for each } b \in DelB: DelB' = DelB' \cup \{ b[\text{source}(id)] \} \\
 Ta'_{Ub} = \langle NewP, DelB', ProcMsg \rangle
 \end{array}$$

**Selection of an Applicable Plan** This transition rule has been modified so that the  $S_{Ap}$  function has two additional parameters: the agent's current affective state  $Ta_\sigma$ , and the agent's rationality level  $ag_{P_r,i}$ . Thus, the plan that the agent selects for its execution, will be influenced by these two parameters. This is another function that can be customized by the programmer; nevertheless, its default implemented mechanism is described in Section 4.5.

$$\begin{array}{c}
 S_{Ap}(T_{Ap}, Ta_\sigma, ag_{P_r,i}) = (p, \theta) \\
 \hline
 \langle ag, C, M, T, Mem, Ta, SelAppl, ast \rangle \rightarrow \langle ag, C, M, T', Mem, Ta, AddIM, ast \rangle \\
 \text{(SelAppl')}
 \end{array}$$

$$\text{where: } T'_\rho = (p, \theta)$$

**Executing an Intention** Following with the notation used in [Vieira *et al.*, 2007],  $i[p]$  denotes an intention  $i$  with the plan  $p$  on top of it. Similarly to other transition rules above, in the next two rules, if the intention to

## 4. Extension of the AgentSpeak's Operational Semantics

---

be executed implies adding or removing a belief<sup>1</sup>, these beliefs are stored in  $Ta_{Ub}$  for them to be processed in the step  $AffModB$ .

$$\frac{T_l = i[head \leftarrow +b; h]}{\langle ag, C, M, T, Mem, Ta, ExecInt, ast \rangle \rightarrow \langle ag, C', M, T, Mem, Ta', AffModB, ast \rangle} \quad (AddBel')$$

$$\begin{aligned} \text{where: } Ta'_{Ub} &= \langle \{b[source(self)]\}, \{\}, ExecInt \rangle \\ C'_I &= (C_I \setminus \{T_l\}) \cup \{i[head \leftarrow h]\} \end{aligned}$$

$$\frac{T_l = i[head \leftarrow -at; h]}{\langle ag, C, M, T, Mem, Ta, ExecInt, ast \rangle \rightarrow \langle ag, C', M, T', Mem, Ta', AffModB, ast \rangle} \quad (DelBel')$$

$$\begin{aligned} \text{where: } Ta'_{Ub} &= \langle \{\}, \{at[source(self)]\}, ExecInt \rangle \\ C'_I &= (C_I \setminus \{T_l\}) \cup \{i[head \leftarrow h]\} \end{aligned}$$

**Affective modulator of beliefs** In this transition, beliefs to be added and removed in tuple  $Ta_{Ub}$  are modulated according to the agent affective state  $Ta_\sigma$  (by the function  $modB(AddB, DelB, Ta_\sigma)$ ), where a new set of beliefs to be added ( $MAddB$ ) and removed ( $MDelB$ ) are obtained. The third component of  $Ta_{Ub}$  indicates the step that requires the addition or deletion of beliefs, helping to determine the next step in the cycle (i.e.,  $SELv$  or  $CLrInt$ ). The corresponding additions and deletions are performed, and the corresponding events of belief addition or deletion are created.

$$\begin{aligned} Ta_{Ub} &= \langle AddB, DelB, ProcMsg \rangle \\ modB(AddB, DelB, Ta_\sigma) &= \langle MAddB, MDelB \rangle \\ \hline \langle ag, C, M, T, Mem, Ta, AffModB, ast \rangle &\rightarrow \langle ag', C', M, T, Mem, Ta, SELv, ast \rangle \\ &\quad (ModB_1) \end{aligned}$$

---

<sup>1</sup>In AgentSpeak the beliefs that are added through an action of a plan are called “mental notes”.

#### 4. FORMALIZATION OF GENIA<sup>3</sup>

---

$$\begin{aligned}
 \text{where for each } mdb \in MDelB: \quad & ag'_{bs} = ag_{bs} \setminus \{mdb\} \\
 & C'_E = C_E \cup \{\langle -mdb, \top \rangle\} \\
 \text{and for each } mab \in MAddB: \quad & ag'_{bs} = ag_{bs} \cup \{mab\} \\
 & C'_E = C_E \cup \{\langle +mab, \top \rangle\}
 \end{aligned}$$

$$\begin{aligned}
 & Ta_{Ub} = \langle AddB, DelB, ExecInt \rangle \\
 & \text{modB}(AddB, DelB, Ta_{\sigma}) = (MAddB, MDelB) \\
 \hline
 & \langle ag, C, M, T, Mem, Ta, AffModB, ast \rangle \rightarrow \langle ag', C', M, T, Mem, Ta, ClrInt, ast \rangle \\
 & \hspace{15em} (\text{ModB}_2)
 \end{aligned}$$

$$\begin{aligned}
 \text{where for each } mdb \in MDelB: \quad & ag'_{bs} = ag_{bs} \setminus \{mdb\} \\
 & C'_E = C_E \cup \{\langle -mdb, \top \rangle\} \\
 \text{and for each } mab \in MAddB: \quad & ag'_{bs} = ag_{bs} \cup \{mab\} \\
 & C'_E = C_E \cup \{\langle +mab, \top \rangle\}
 \end{aligned}$$

**Mood temporal dynamic** A single cycle controls how the affective state decays over time. It contains the single step `AsDecay`, which is continuously executed. This task is performed by the `AsDec` ( $Ta_{\sigma}, ag_{P_{tr}}$ ) function.

$$\begin{aligned}
 \hline
 & \langle ag, C, M, T, Mem, Ta, s, ast \rangle \rightarrow \langle ag, C, M, T, Mem, Ta', s, ast \rangle \\
 & \hspace{15em} (\text{DecAffState})
 \end{aligned}$$

$$\text{where: } Ta'_{\sigma} = \text{AsDec}(Ta_{\sigma}, ag_{P_{tr}})$$

The transition rules that correspond to the steps not previously addressed have also been modified so that the structure of the agent configuration has been adapted to the new configuration. They are not presented for simplicity.



## 4.5 GenIA<sup>3</sup>'s Default Design

In Section 4.4, a set of functions that are used in the agent execution cycles, and that can be customized by the programmer were introduced. The existence of this set of functions adds flexibility to the agent programmer to adapt the agents' behavior to different psychological theories and application domains. In this section we offer a default implementation of these functions, which is based on widely used psychological theories. This default implementation has been used in several scenarios, and can be extended if required. Next, it is performed a general description of these functions, in order to offer a global understanding of the *GenIA*<sup>3</sup>'s default implementation, and avoiding too specific details for simplicity. Alternative designs can be found in [Alfonso *et al.*, 2016b] and in Chapters 5 and 6.

We propose a design for the steps of the affective cycle which is inspired by the ALMA model [Gebhard, 2005] and the EMA appraisal model [Marsella & Gratch, 2009]<sup>1</sup>. The components and processes of the default design include:

**Affective state:** In *GenIA*<sup>3</sup>'s default design, the affective state  $Ta_\sigma$  is represented as the agent's mood in a dimensional way, where three values in a range  $[-1, 1]$  describe the agent's mood in a given time. These values correspond to pleasure, arousal, and dominance (or PAD, according to Mehrabian's model [Mehrabian, 1996b] of mood).

**Appraisal variables:** Appraisal variables  $Ta_{Av}$  include: desirability, likelihood, expectedness, controllability, and causal attribution. These are appraisal variables selected from the EMA model, and their values are in a range  $[0, 1]$ , with the exception of expectedness which is in a range  $[-1, 1]$ .

**Personality traits:** The traits of the agent's personality follow the Five Factor Model [McCrae & John, 1992], which describes individual traits

---

<sup>1</sup>In the implementation of the steps we avoided to introduce too much execution complexity, selecting as default mechanisms those most commonly implemented in computational approaches.

## 4. FORMALIZATION OF GENIA<sup>3</sup>

---

through five dimensions (openness, conscientiousness, extraversion, agreeableness, and neuroticism) in a range  $[0, 1]$ .

**Affective state initialization:** The initial (and also equilibrium affective state) of the agent is calculated following Mehrabian’s work [Mehrabian, 1996a], which proposes a mapping of the agent’s five dimensions of personality to the three dimensions of the PAD space.

**Appraisal process:** As part of this process, the function `Appraise` ( $\varepsilon, bs, cc, Mem, Ap$ ) evaluates the event  $\varepsilon$  when this event implies the *addition* or *deletion* of a belief. This function determines the values of the appraisal variables on the basis of the resulting state, and after the addition or deletion of the belief. The values for the appraisal variables,  $Av$ , are calculated as follows:

- **Desirability:** is determined according to the agent’s concerns (by using its numerical value).
- **Likelihood:** is always one in the default implementation, since it is assumed that there is no uncertainty about the state of the world.
- **Expectedness:** is determined according to the time range of the ‘belief probability’. Lets assume that the event  $\varepsilon$  corresponds to the addition of a belief  $b$  at time  $t^1$ , and that  $b$  is represented with a probability  $p$  and a time range  $R$ . If  $t$  is in  $R$  the expectedness of  $\varepsilon$  is the value of  $p$ , and if  $t$  is not in  $R$  the expectedness of  $\varepsilon$  is the negative value of  $p$ . Besides, if  $p$  is zero (meaning that the agent doesn’t expects  $b$  at all), then the expectedness takes the value of  $-1$  if  $t$  is in  $R$  and  $1$  otherwise<sup>2</sup>.
- **Controllability:** Considering that it expresses whether the event outcome can be altered in order to block some effect, in the

---

<sup>1</sup>Milliseconds since the system starts its execution.

<sup>2</sup>Although not used in the default design, *GenIA*<sup>3</sup>’s platform has tools for determining fulfilled and not fulfilled expectations in a more comprehensive way. See [Taverner *et al.*, 2016] for more details.

default design it is the difference between the desirability after  $e$  has an effect in the agent cognition and the desirability before the  $e$ 's effect.

- **Causal attribution:** It can be the environment (if the belief to be added or removed is a percept), other agent (if the belief to be added or removed is a message), or self (if the belief to be added or removed is a mental note).

**Affectively relevant events:** In the default implementation, the function  $\text{AffRelEv}(\varepsilon, Av)$  determines that the event  $\varepsilon$  is relevant when the desirability in  $Av$  is not in a range of “average desirabilities” (i.e., when it is extremely undesirable or extremely desirable).

**Affective state updating process:** The affective state is updated through the function  $\text{UpAffSt}(\sigma, Av)$  in three steps. Firstly, five possible emotion categories can be derived (surprise, joy, sadness, anger, and guilt), starting from the appraisal variables used in [Marsella & Gratch, 2009]. Secondly, each emotion is mapped into the three PAD dimensions following [Gebhard, 2005]. Thirdly, mapped emotions are averaged in a single value for each dimension according to [Gebhard, 2005]<sup>1</sup>. Finally, the current affective state is updated by using the calculated affective state through the “pull and push” mood change function proposed in [Gebhard, 2005]. This function can be parametrized regarding the magnitude in which the current affective state is ‘moved’ from its current position.

**Selection of actions:** The function  $S_{Ap}(Ap, \sigma, P_{rl})$  uses the affective state  $\sigma$  and the rationality level of the agent personality  $P_{rl}$  (rat\_level in the grammar of Figure 4.2) in order to select the next actions to be performed (by selecting the next applicable plan). This function selects an applicable plan by ranking applicable plans with and without

---

<sup>1</sup>Considering that the set of emotions in [Gebhard, 2005] doesn't contain all emotions in [Marsella & Gratch, 2009] we carefully looked for a similarity, assuming sadness as distress and guilt as remorse.



```

1 MAS example_default_MAS {
2   infrastructure: Centralised
3   agents:
4     player
5       agentClass jason.Em.asSemantics.EmAgent
6       agentArchClass jason.Em.architecture.EmAgArch
7       emEngineClass jason.Em.architecture.DefaultEmotionEngine
8       personalityClass jason.Em.architecture.OCEAN
9       affStateClass jason.Em.architecture.PAD;
10 }

```

Figure 4.5: Code of the MAS configuration file for an agent ‘player’ following the default design.

*Now*,  $Now + DayMsecs >$ ), and that his previous score in this class was 3. If he is told by the teacher that he passed with 4, desirability will be  $\frac{4}{5}$ , likelihood will be 1, and the causal attribution will be other agent. Besides, Expectedness will be 0 and controllability will be  $\frac{4}{5} - \frac{3}{5}$  (i.e., the value of his concerns after he knows the score and the value of his concerns before he knows the score). With these values of the appraisal variables the emotion joy is triggered.

An agent following the default design of *GenIA*<sup>3</sup> should be declared in the MAS configuration as shown in Figure 4.5. This figure shows a MAS with a single agent “player”. The agent’s class and architecture’s class must be *EmAgent* and *EmAgArch* respectively. Besides, the default classes for the engine that manages the agent’s affective processes, the agent’s personality, and the agent’s affective state should be *DefaultEmotionEngine*, *OCEAN*, and *PAD* respectively. These classes can be extended in order to define a specialized version of *GenIA*<sup>3</sup>. Additionally, the infrastructure that should be used in order to work with *GenIA*<sup>3</sup>’s affective agents is *Centralized*. This infrastructure is an extended version of the Jason’s centralized architecture (version 2.0.0 of Jason). Examples of the default design and of extended designs can be found in Chapters 5 and 6.

### 4.6 Conclusions

This chapter has offered a formalization of *GenIA*<sup>3</sup>, a General-purpose Intelligent Affective Agent Architecture which is based on the BDI agent architecture. This formalization includes an extension of the AgentSpeak reasoning cycle, the definition of its operational semantics, and the extension of the syntax of an AgentSpeak-based agent language (Jason) to include affect-related attributes, thus achieving the Goals G2.2 and G2.3 of this thesis.

With this kind of formalization, approaches can be adapted to specific application domains' requirements and psychological theories, and comparisons of different psychological theories can be easily performed. More flexibility is also provided to the language, by offering a compact and elegant notation that avoids ambiguities. *GenIA*<sup>3</sup> is grounded on widely studied psychological and neurological theories and offers an integral vision of the agent and its behavior, considering both rational and affective attributes and processes. In order to offer this integral vision, we've not only modeled emotions. We have also addressed the agent affective state in a more generic way, allowing the use of different psychological theories for its representation. Also, different psychological theories can be used in other affect-related processes, since our formalization allows customizing several steps in the reasoning and affective cycles. This allows the creation of agents of diverse nature such as agents whose affect influences their cognition, or virtual characters able to express emotions.

On the other hand, when defining the agent personality, it is possible to include other characteristics besides the personality traits, traditionally addressed by computational approaches. In our approach, it is also possible to define a "rationality level" for the agent, as well as different frequencies for its affective and reasoning cycles. This fulfills Goal G2.5, allowing to define whether an agent is more or less 'emotive'. In the proposed architecture the agent's response tendencies (or coping strategies) for a given situation and affective state can also be defined. Thus, the affective state not only drives goals, but also may generate new ones through the agent coping strategies. We also provide a default implementation, thus fulfilling Goal G2.4. With this

default design any expert in fields like psychology or behavioral computing, could use it in order to provide more precise and refined ways of describing each particular affective process. Chapters 5 and 6, describe some case studies where extended versions of the *GenIA*<sup>3</sup>'s are proposed. Moreover, in Chapter 6, a comparison between solutions with the default design and with the extended design is also performed. These solutions have been implemented by using the extension of the Jason's MAS platform described in this chapter, and which fulfills the Goal G3 of this thesis.





**Part IV**

**Evaluation**



## Introduction

As exposed in Chapter 1 the main motivation of this thesis has been the simulation of believable human behaviors and a believable human-like decision making process. This motivation arose as part of the *mWater* project [Alfonso *et al.*, 2014a; Giret *et al.*, 2011], whose main goal was the development of a virtual market for trading water rights. In order to simulate humans' affective state and its influence on humans' decision-making through *GenIA*<sup>3</sup>'s agents, example solutions of the classical games used in behavioral economics for addressing social preferences were designed. These solutions are described later on Chapter 5. Besides, in Chapter 6, the *GenIA*<sup>3</sup>'s capabilities to simulate human behavior in real situations are validated. To this end, we designed and performed an experiment with real people whose mood was induced, and whose decisions, in a Black Jack game, were recorded. With this experiment, it was determined how mood changes for a given event, and how mood influences subsequent decisions. This experiment corresponds to the Goal G4 of this thesis.

The case studies have been conceived by always taking into account ethical considerations (see Appendix A.1 for a related discussion). For all case studies we adapted *GenIA*<sup>3</sup> so that the appraisal, the update of the affective state, and the decision-making processes were improved. In particular, Chapters 5 and 6 aim to achieve the Goal G5 of this thesis and its subgoals. Specifically, in these chapters, it is aimed to:

- Verify that *GenIA*<sup>3</sup>'s representation of affect and personality allows to define agents' behavior more in line with human behavior than traditional methods (Goal G5.1).
- Validate *GenIA*<sup>3</sup> by verifying that it is able to represent situations where (Goal G5.1):
  - There are different behaviors for different affective states.
  - There are different behaviors for different personalities, including different rationality levels.

## EVALUATION

---

- Verify that the default design of *GenIA*<sup>3</sup> is able to be adapted to fulfill specific requirements (Goal G5.2).
- Verify that, by adapting the default design to use alternative psychological theories, results can be improved (Goal G5.2).
- Verify that affective agents represent human behavior better than merely rational agents (Goal G5.3).

---

---

## Affect and Personality in Classical Games for Addressing Social Preferences

### 5.1 Introduction

Classical games of game theory are an important tool in experimental economics to evaluate how individuals behave in situations where a single decision must be taken [Swope *et al.*, 2008]. These games often address social preferences, which are one type of preference investigated in behavioral economics that “assumes people are self-interested, but are also concerned about the payoffs of others” [Charness & Rabin, 2002]. Social preferences are related to concepts such as reciprocity, altruism, inequity aversion, or fairness. Examples of these classical games are the ‘Trust game’, the ‘Ultimatum game’, the ‘Dictator game’ and the ‘Prisoner’s dilemma’. In many of these games there are generally two players involved, and there is also a stable equilibrium concept called “Nash equilibrium”. This equilibrium is a concept that represents the set of players’ decisions where no one can obtain more benefits by making a different decision while the other players’ decisions keep unchanged [Nash, 1951]. Nevertheless, what happens in real situations is that individuals take decisions biased by many external or individual factors. It has been proved [Adolphs, 2003; Olsson &

## 5. AFFECT AND PERSONALITY IN CLASSICAL GAMES

---

Ochsner, 2008], that the affective characteristics are one of the most influential factors on social decision making. We have selected the Prisoner's Dilemma and the Trust Game in order to demonstrate how *GenIA*<sup>3</sup>'s agents can simulate human decisions influenced by their affect and/or personality, and how these affective agents also solve the complex problems that rational agents solve [Alfonso *et al.*, 2015b].

### 5.2 Prisoner's Dilemma

The Prisoner's Dilemma (PD) has been a useful tool for many disciplines, e.g. economics, sociology, politics, etc. [Sewell, 2010]. It offers a model of cooperation that emphasizes how individual and collective interests coexist [Axelrod & Hamilton, 1981]. The participants represent two prisoners. Each one has the opportunity to declare that the opponent committed the crime and hence to **betray** him, or otherwise **cooperate** with the opponent and keep silence. Each possible combination of decisions has a punishment of a number of years in prison for the participants. If prisoner *A* and prisoner *B* cooperate each other and they keep silence, then they both receive a year in prison as punishment; if one betrays and the opponent keeps silence, then the first receives the temptation punishment of zero years (i.e., no punishment at all), and the opponent receives the sucker punishment of three years in prison; if both prisoners betray each other, then they both receive the defect punishment of two years in prison [Tucker, 1983]. We used the iterated version of the prisoner's dilemma, where there are  $N$  rounds of decisions before the final decision. Playing the PD more than once helps to observe how the participants make decisions under the influence of the opponent's previous actions.

#### 5.2.1 Rational vs. Human Behavior in the Prisoner's Dilemma

The economic theory predicts the third option (prisoners betray each other) as the most rational if each player evaluates his own possibilities independently of what the opponent decides, and assuming that the players never trust each other. When mutual trust is the case, then the dilemma ceases to be a dilemma

## 5.2. Prisoner's Dilemma

	2	0	0	2	2	2	2	2	2	2	2	2	2	2	Years
betraying agent	#7@*	#7@*	#7@*	#7@*	#7@*	#7@*	#7@*	#7@*	#7@*	#7@*	...	#7@*	196		
TESTER agent	#7@*	+zZ	+zZ	#7@*	#7@*	#7@*	#7@*	#7@*	#7@*	#7@*	...	#7@*	202		
	2	3	3	2	2	2	2	2	2	2	2	2			

Figure 5.1: Plays of the *betraying* agent against the *TESTER* agent in 100 rounds of the prisoner's dilemma.

[Sewell, 2010] and the best solution would be that both players continuously cooperate each other (keeping silence). In an iterated version of the prisoner's dilemma with rational agents, that act selfishly and that never trust each other, agents would always select the option of betraying. Nevertheless, results of experiments systematically demonstrate that when playing the iterated PD, humans do not follow what the economic theory predicts is the best strategy. They tend to show aims of cooperation instead (i.e., they tend to keep silence to a higher or lesser degree) [Clark & Sefton, 2001; Khadjavi & Lange, 2013; Ortmann & Tichy, 1999].

Many strategies could be used for the iterated PD that consider the opponent's decisions. In 1979 Robert Axelrod held a tournament that aimed to rank participants' strategies for playing 200 rounds of the PD against each other strategy [Axelrod, 1984]. The winner was a simple strategy called Tit-For-Tat (TFT), which consisted on cooperating always in the first round, and then reproducing, on each of the next rounds, what the opponent did on the previous round. Later, Axelrod held another tournament with some variations, where more competing strategies participated. Again TFT won the tournament, even when participants were aware of its characteristics, and also aware of the analysis of the results of the first tournament. According to Axelrod [Axelrod, 1984], the success of TFT in computer tournaments has several reasons. Some of them include: 1) TFT is *nice*, because it is never the first to betray; 2) TFT is *retaliatory*, because it responds to the same extent to the opponent's actions; 3) TFT is *forgiving*, because it is able to cooperate again no matter what were the opponent's previous behavior; and 4) TFT is *clear*, because the opponent can easily understand it, what favors long-term cooperation.

Figure 5.1 shows the results of our implementation of an agent that always

## 5. AFFECT AND PERSONALITY IN CLASSICAL GAMES

---

betrays (hereinafter *betraying* agent), playing 100 rounds against an agent that plays the TESTER strategy (hereinafter *TESTER* agent). In the TESTER strategy, the player starts betraying. The aim of this first move is to test the opponent's response. In the next move the player cooperates, and the opponent's next move determines whether the *TESTER* agent changes the initial strategy. If the opponent answers to a betrayal by betraying, then the player apologizes by cooperating again and, in the next moves, the player just reproduces the movements of the opponent. If the opponent cooperates after being betrayed, the TESTER takes advantage and betrays again in the next move. In this example, always betraying is the best strategy against a TESTER opponent. Nevertheless, if agents would consider the opponent's previous decisions in an iterated PD, then always betraying can have consequences, because the opponents could take advantage of it. For example, consider including the forgiving Tit-For-Tat (FTFT) strategy. It reproduces the opponent's moves but it requires two betrayals before it betrays. Figure 5.2 shows the results of our implementation of three agents (*rational*, *TESTER*, and *FTFT*) playing 100 rounds against each other. In this example, if we consider the total amount of years collected, the *TESTER* agent performs much better than the *rational* and the *FTFT* agents, because the *TESTER* agent accumulates only 252 years compared to the 392 years of the *rational* agent and the 402 years of the *FTFT* agent. This shows that the rational strategy is not always the best.

### 5.2.2 GenIA<sup>3</sup> Agents' Behavior in the Prisoner's Dilemma

In [Sewell, 2010] Sewell suggests that emotions support humans when solving the PD, by enabling cooperation. The author makes an analogy of the TFT's characteristics with human emotions, by comparing 1) being *nice* with a feeling of *well-being* after mutual cooperation, and with *gratitude*, which fosters cooperation; 2) being *forgiving* with *guilt*, which is what may be experienced after receiving cooperation from the opponent in response to a betrayal, which also leads to cooperate in the next move; 3) and being *retaliatory* with *anger* after being betrayed by the opponent, which leads to a betrayal in the next move.



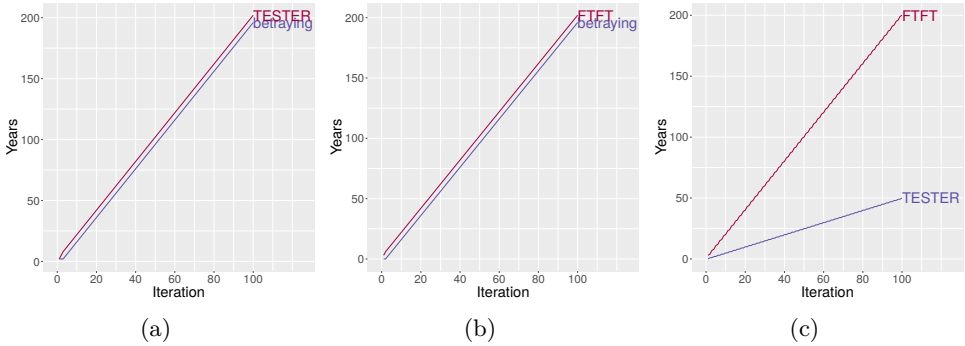


Figure 5.2: Comparison of the years in prison of the *betraying* agent, the *FTFT* agent, and the *TESTER* agent playing each other 100 rounds of the prisoner's dilemma.

We have performed a simulation in the *GenIA*<sup>3</sup>'s platform by using these analogies. In this simulation an affective player (*playerAff*) makes the same plays of a *TFT* agent, but it has an affective state whose changes depend on the opponent's plays. Also *playerAff*'s decisions depend on its affective state. Figure 5.3 shows an extract of the output of the *playerAff* agent playing 100 rounds of the PD against the *TESTER* agent. An extract of the code of the *GenIA*<sup>3</sup> agent *playerAff*, as well as of the *TESTER*, *FTFT*, and *betraying* agents can be found in Appendix B.4. *playerAff*'s personality is defined through the trait 'cooperativeness' (which can have two values: 1 for cooperative and 0 for non cooperative), and through the rationality level (which can have two values: 1.0 for completely rational and 0.0 for completely emotive). In this simulation *playerAff*'s personality is defined as follows:

```
personality__ : { [0], 0.0 }.
```

It is a non-cooperative and completely emotive personality. The non-cooperative personality defines its initial affective state, which is 'angry'. Consequently, its initial play is to betray. The zero value rationality level indicates that he behaves in a completely emotional way. According to *playerAff*'s behavior, when it is arrested it will 'betray' when feeling angry, 'cooperate' when feeling

## 5. AFFECT AND PERSONALITY IN CLASSICAL GAMES

---

```
1 [arbiter] I'm arresting playerAff and TESTER (time 0)
2 [TESTER] I played b against playerAff at time 0
3 [playerAff] I played b(angry) against TESTER at time 0
4 [TESTER] Got 2 at time 0. My total now is: 2
5 [playerAff] Got 2 at time 0. My total now is: 2
6 [arbiter] I'm arresting playerAff and TESTER (time 1)
7 [TESTER] I played c against playerAff at time 1
8 [playerAff] I played b(angry) against TESTER at time 1
9 [TESTER] Got 3 at time 1. My total now is: 5
10 [playerAff] Got 0 at time 1. My total now is: 2
11 [arbiter] I'm arresting playerAff and TESTER (time 2)
12 [TESTER] I played c against playerAff at time 2
13 [playerAff] I played c(guilty) against TESTER at time 2
14 [TESTER] Got 1 at time 2. My total now is: 6
15 [playerAff] Got 1 at time 2. My total now is: 3
16 [arbiter] I'm arresting playerAff and TESTER (time 3)
17 [TESTER] I played c against playerAff at time 3
18 [playerAff] I played c(grateful) against TESTER at time 3
19 ...
```

Figure 5.3: Extract of the output of 100 rounds of the PD played between the *playerAff* and the *TESTER* agents.

‘grateful’ or ‘guilty’, and ‘betray’ by default (i.e., when there is no influence of the affective state). Figure 5.3 shows that when *playerAff* confirms that its opponent has betrayed him in the previous round (time 0), he gets angry and he decides to betray in the next round. In the third round he feels guilty after being informed of the cooperation of his opponent in the second round, so that he also cooperates. In the fourth round he also cooperates because he feels grateful due to the cooperation of his opponent in the third round.

We have extended the example of the prisoner’s dilemma included in Jason release 2.0<sup>1</sup> [Bordini *et al.*, 2007] by including the above mentioned players *playerAff*, *TESTER*, and *FTFT*. An example of the code of the MAS configuration file is shown in Figure B.4.1. The MAS simulation is performed with these agents: *arbiter*, *TESTER*, and *playerAff*. Specifically the agent *playerAff* is defined as follows:

```
playerAff
  agentClass jason.Em.asSemantics.EmAgent
```

---

<sup>1</sup><https://sourceforge.net/projects/jason/>

```

agentArchClass jason.Em.architecture.EmAgArch
emEngineClass pdclasses.pd_emEngine
personalityClass pdclasses.pd_personality
affStateClass pdclasses.pd_affect;

```

Considering that the agent *playerAff* is an affective agent, the classes *EmAgent* and *EmAgArch* should be used for his definition and for the definition of his architecture respectively. Additionally, three specialized classes were defined for managing *playerAff*'s affective cycle (*pd\_emEngine*), *playerAff*'s personality (*pd\_personality*), and *playerAff*'s affect (*pd\_affect*). The class *pd\_emEngine* updates the affective state with the value of the agent concerns when an event of the kind *last\_move* is received. This means that, when the agent is aware of the opponent's last move, his affective state may be affected. *playerAff*'s affective state is defined through a single trait 'pleasure' (which can have a value between 0.0 and 1.0). Also, in the class *pd\_emEngine*, the agent's initial affective state is set according to the agent personality in such a way that, if the agent is cooperative, his initial affective state is 'grateful' and, if the agent is non-cooperative, his initial affective state is 'angry'. These mood labels are defined in the MAS configuration file as follows:

```

affcategories: grateful (<0.68:1.0>),
                 guilty (<0.35:0.67>),
                 angry (<0.0:0.34>);

```

On the other hand The extent to which *playerAff* opponent's and self moves affect *playerAff*'s concerns, is expressed in the agent's concerns:

```

concern1 (C1) :- last_move (O, OLM) &
                 ( (OLM==c & C1=2) | (OLM==b & C1=0) ).
concern2 (C2) :- my_last_move (MLM) &
                 ( (MLM==c & C2=1) | (MLM==b & C2=0) ).
concerns__ (V) :- concern1 (C1) & concern2 (C2) & V=(C1+C2)/3.

```

When the self or the opponent's last move is to betray, the concerns' value is low (zero in this case), what means that this is not a nice move for the agent.

## 5. AFFECT AND PERSONALITY IN CLASSICAL GAMES

---

On the other hand, if the self or the opponent's last move is to cooperate, then this is nice for the agent. Similarly it is nicer when the 'cooperation' last move is from the opponent (first concern's value of 2), than when the 'cooperation' last move is self (second concern's value of 1). The final value of concerns is the sum of these two individual concerns divided by 3.

Note that an affective agent that has a non-cooperative personality, and that is completely rational (i.e., its rationality level is 1.0) will perform exactly the same than an agent that always betrays (no matter which are the movements of the opponent). This agent would get a better compensation of years in prison than the *TESTER* agent. Additionally, by transforming *playerAff* into a cooperative agent (by assigning 0.0 to the personality's rationality level), at the beginning it would feel in a 'grateful' mood, but this mood would change immediately in the second round after being informed of the betrayal of the *TESTER* agent in the first round. Thus, by being cooperative, *playerAff* would get the same compensation in years than *TESTER*. In summary, the affective agent has a behavior identical to that of an agent playing TFT, because the simplicity of the problem and the few possible decisions doesn't allow for more complex designs. Nevertheless, the way of addressing the problem exposed in this example offers a wide amount of possibilities for adding a human component to the PD players. For example, more affective categories can be defined, new personality traits and values for these traits can be included, or more complex concerns could be defined.

### 5.3 Trust Game

Another important classical game used in experimental economics is the Trust Game (TG) [Berg *et al.*, 1995]. In this game a proposer (or *trustor*) is given an amount of endowment to be split and shared with a *trustee*. The amount shared  $x$  is multiplied by a factor (usually three), in such a way that, after receiving the endowment, the *trustee* has an amount of  $3 * x$  available. Then, the *trustee* must decide to send back any amount to the *trustor*. This way the *trustor* hopes to receive something back related to its initial offer, trusting in the intentions of the other.

### 5.3.1 Rational vs. Human Behavior in the Trust Game

In the Trust Game (TG) the perfect equilibrium is obtained by not trusting at all [Bracht & Feltovich, 2008]. Nevertheless, in real situations of the TG both participants tend to share more than what the perfect equilibrium predicts. For example in [Harth & Regner, 2016], experiments with human participants show that in an iterated TG, when the trustor receives feedback about whether the trustee gives something in return (and how much), the information about back transfers triggers the emotion ‘anger’ on trustees, and hence, they are less willing to cooperate. This study also shows that, personality’s characteristics also influence the level of anger triggered in trustees. Specifically, the ‘negative reciprocity orientation’ (NRO), which is “the tendency to return unfavorable treatment” [Harth & Regner, 2016], appears to have a direct relation with the levels of anger and hence, an inverse relation with the willingness of cooperation. In order to reproduce real circumstances in the TG, we used this study (hereinafter the ‘base study’). We simulate how anger is triggered, how anger changes throughout the rounds, and how it influences the amount sent by the *trustor* in situations where a feedback is received.

### 5.3.2 GenIA<sup>3</sup> Agents’ Behavior in the Trust Game

The simulation was performed through a MAS, where 6 rounds of the TG were played<sup>1</sup>. Three agents were part of the MAS: *moderator*, *trustor*, and *trustee*. The *trustor* was designed as an affective agent, and the *trustee* simply played a random play on each round. The code of the *trustor* is shown in Appendix B.5. Taking into account the first step of the stepwise regression performed in the base study [Harth & Regner, 2016], anger is directly related to the difference between the back transfer and the amount sent in the previous round (with a coefficient of  $-0.28$ ). Also, the round of the game is inversely related to the agents’ anger in such a way that, as the round of the game increases, participant’s anger also increases by an amount of 0.24. The *trustor*’s concerns have the following structure:

---

<sup>1</sup>This is the amount of rounds used in the base study.

## 5. AFFECT AND PERSONALITY IN CLASSICAL GAMES

---

**concerns\_\_**(V) :- last\_difference(Round, LastDiff) &  
V = 0.28\*LastDiff-0.24\*Round.

The *trustor*'s concerns were defined as the inverse of 'anger', thus the lesser the value of anger is, the better for the agent. That's why the coefficients in the agent's concerns have the opposite sign of the sign of the corresponding values in the base study. There is another factor that is directly related to the agent's anger after receiving a back transfer, which is the agent personality trait NRO. The *trustor*'s personality has two components that range from 1 to 7, representing the two items from [Dohmen *et al.*, 2009] ("If somebody puts me in a difficult position, I will do the same to him/her", and "If I suffer a serious wrong, I will take revenge as soon as possible, no matter what the cost"). The agent's personality is defined as follows:

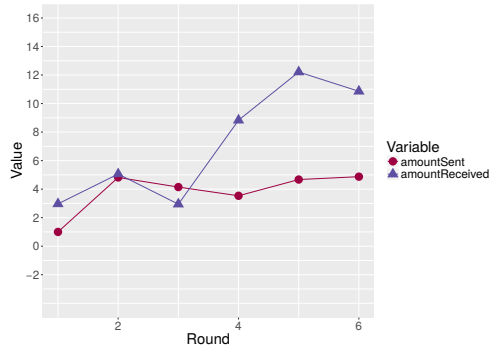
**personality\_\_** : { [6,7], 0.0 }.

The agent's personality is used by the emotion engine when updating the agent's affective state. The affective state, in turn, is a combination of the calculated anger and the agent personality. A higher value of the affective state is related to a higher anger and vice versa.

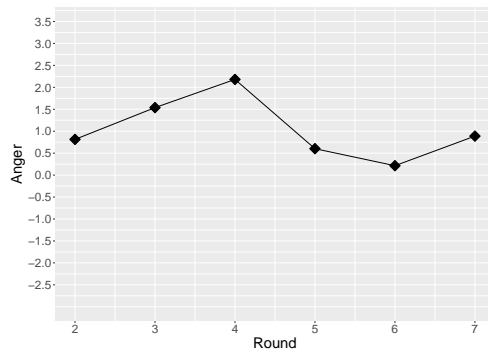
Unlike the PD, in the TG the agent doesn't have different plans for different affective states. Considering that there is a direct relation between the affective state and the next amount to send, there is a single 'affective' plan for a single affective category ('singleCat') that uses the value of the affective state for calculating the next amount to send, which is determined as follows:

$AmToSend = 0.24*LastDiff + AffState - 0.31*PrevRound + 5.98;$

Figure 5.4 shows the results of the amount sent by the trustor, the amount sent back by the trustee, and the level of anger of the trustor for each one of the six rounds. As shown in Figure 5.4(b), for a given round, the level of anger decreases when the difference between the back transfer and the amount sent on the previous round in Figure 5.4(a) is high. Note that the amount sent in the next round also increases. On the other hand, when this difference is small



(a)



(b)

Figure 5.4: Simulation of the agent *trustor* in six rounds of the iterated TG with random back transfers.

## 5. AFFECT AND PERSONALITY IN CLASSICAL GAMES

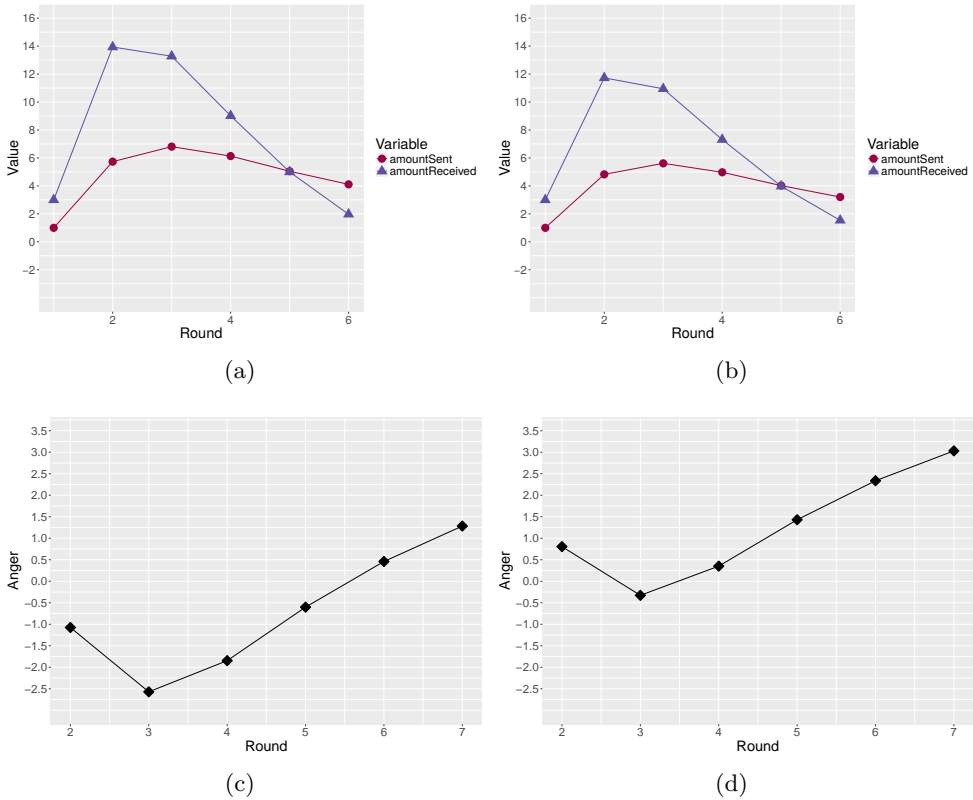


Figure 5.5: Simulation of the agent *trustor* in six rounds of the iterated TG with descending back transfers. (a) and (c) *trustor*'s NRO=[2,3]. (b) and (d) *trustor*'s NRO=[6,7].



or negative the level of anger increases and the amount sent in the next round decreases. In this simulation the initial value of anger was set to zero.

In order to compare the results of two agents with different personalities we designed a *trustee* agent that starts sending high back transfers (relative to the amount sent by the *trustor*) and then decreases the relative amount sent back on each round. A *trustor* with NRO (2,3) (Figures 5.5(a) and 5.5(c)), and a *trustor* with NRO (6,7) (Figures 5.5(b) and 5.5(d)) were compared. Both results are similar in that anger increases with low values of the difference between the back transfer and the amount sent. Nevertheless, the agent with lower NRO has lower levels of anger<sup>1</sup>, and hence, it sends slightly higher amounts of money on each round. It worth mentioning that, when modifying the *trustor*'s rationality level from completely emotional to completely rational (by changing the rationality level from 0.0 to 1.0), the agent always performs the most rational behavior (no trust) and always sends the minimum amount of money allowed through a plan that has the following structure:

```
@psend[atomic]
+newRound(Round, Opponent) [source (moderator) ]
  : minimumPossible(MP) & initialEndowment(IE)
  <- +amountSent(Round, MP);
  AmountLeft = IE-MP;
  +amountLeft(Round, AmountLeft);
  .send(moderator, tell, play(Round, MP));
  .print("I send ", MP, " to the trustee at time ", Round).
```

---

<sup>1</sup>In the examples of this Section anger can have negative values because the data set used is different from the data set of the base study. Nevertheless this is not a relevant issue since our interest is to show that the tendency of the data is correct and not to offer final results.



---

## GenIA<sup>3</sup>'s Agents Playing Black Jack. A Case Study

One of the bottlenecks of research in mood and emotion modeling has always been the scarcity of data in order to validate and adjust architectures and platforms. In order to gather this data, we have designed and carry out an experiment where participants' mood variations, and the impact of mood on decisions were addressed in the context of a version of the Black Jack game. In this kind of experiments the mood and emotion measurement is often a challenging task where several techniques can be used. Affective characteristics can be measured as *feelings* in terms of activation and arousal (Ej., skin resistance, muscle tension, cardiovascular measures), as *bodily reactions* in terms of valence and arousal (Ej., heart rate, sweat gland, pupil diameter), or as *emotional expressions* (Ej., by using a coding system such as the Facial Action Coding System [Ekman & Friesen, 1978], or by electromyography). Nevertheless, in experiments of mood and emotion induction, one of the most commonly used measurements of affective characteristics is the self-report. Although self-reports have problems (such as the inaccuracy of the reports of the emotional states), they can be very useful and valuable, specially when other measures are also part of the experiments [Coppin & Sander, 2016]. They

## 6. GENIA<sup>3</sup>'S AGENTS PLAYING BLACK JACK

---

allow performing an easy data collection without needing a high technical support. For this reason we have used this technique in the experiment that has been designed. We have included mood measurement at the beginning, at the middle, and at the end of the experiment. This section describes the characteristics of the Black Jack game, the design of the experiment, data analysis, and the results of the simulation of the experiment through *GenIA*<sup>3</sup> agents.

### 6.1 Black Jack Experiment

#### 6.1.1 General Description

In this experiment a version of the Black Jack game described in [Griffin, 1996] was used. In our version of the game there is a single player who plays against the bank. For each game there is a single complete deck at the beginning of each game, and the cards are represented in black and white with the corresponding value drawn on it. The representation was not performed with original cards and the original name of the game was never used, avoiding any kind of subjective influence in participants familiarized with the game (see Figure 6.1).

The rules of the game can be described as follows. The player's goal is to either surpass the amount of points of the bank's cards or to reach 21 points with the player's cards. At the beginning of the game both, the player and the bank, are given two cards, and one of the cards of the bank is hidden. The player must bet before the cards are dealt. If, after betting, the player reaches 21 points with his cards, he wins the game. This also holds for the bank, if it reaches 21 points either at the beginning of the game, or when it is on its turn to play. If the player doesn't reach 21 points, he can choose between doubling the bet, hitting a new card or standing. In case of doubling, he gets a new card and its turn to play finishes automatically. When he finishes his turn, the bank takes the turn and it will remain hitting cards until it reaches 17 or more points.

The experiment has been designed in order to mainly answer some questions:

GTIA  
Grupo de Tecnología Informática  
Inteligencia Artificial

UNIVERSITAT POLITÈCNICA DE VALÈNCIA  
DSIC

Partida: 3

De €0.1 a €4.0, ¿cuánto deseas apostar en la próxima partida?:

€ 3,0

Apostar

Tu apuesta final: €3.0

Tu histórico de ganancias por partida

Tus cartas: 2 10 10

Cartas banca: 10

Doblar

Pedir

Plantarse

Puntos usuario: 12

Puntos banca:

REGLAS DEL JUEGO

Figure 6.1: Snapshot of the 3<sup>rd</sup> game of the Black Jack experiment. The player has bet and he must decide between doubling, hitting, or standing.

## 6. GENIA<sup>3</sup>'S AGENTS PLAYING BLACK JACK

---

1) how an individual's mood changes after a single game?, 2) how mood affects decisions during the game? is there a difference on the decisions taken by individuals in different moods?, and 3) is there a difference between individuals in a positive mood and individuals in a negative mood regarding the extent to which mood changes? For answering these questions we started by demonstrating the next hypotheses:

**Hypothesis 1** *The mood after a won or lost game is significantly different from the mood before the game.*

**Hypothesis 2** *The extent to which mood changes after a single game doesn't depend on the mood induction (positive, negative, or none).*

**Hypothesis 3** *Agents under the same circumstances (game), and with a different mood, make different decisions.*

### 6.1.2 Design of the Experiment

For a better comprehension of the characteristics of the experiment and the circumstances on which it has been performed, a description of its design is offered in this section. The characteristics of the participants, the variables assessed, and the structure and characteristics of the Black Jack games played by the participants are described.

Three hundred participants, most of them students or staff of the Universitat de València participated on the experiment in exchange of a monetary compensation (a mean of €12). Specifically, they were said that they would be compensated according to what they could accumulate in five randomly selected games (possible bets ranged from €0.1 to €4.0). If the accumulated value was less than €5.0, they would receive €5.0. Besides, at the end of the experiment, the participant who finished in the first position of the ranking would receive €40.0. The sample included 159 female participants (53%) and 141 male participants (47%). The experiment took place in the LINEEX laboratories of the Universitat de València. The participants' mean age was of 22.11 years (SD = 6.67). The experiment was performed in 5 different sessions

## 6.1. Black Jack Experiment

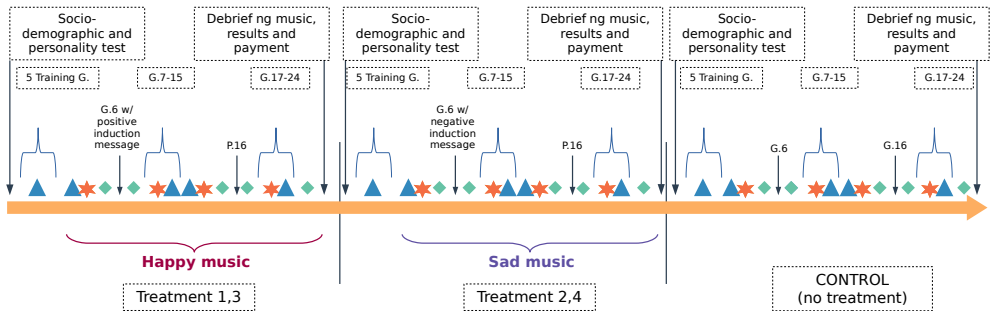


Figure 6.2: Timeline of the Black Jack experiment by kind of treatment. Games can be preceded by an “aspiration” (triangles), an “autoefficacy” (stars), and/or a mood test (diamonds).

of 60 participants each. The dependent variables were the mood variations after a game and the decisions taken in a game.

In order to address the hypotheses of Section 6.1.1, participants underwent five kind of game situations or treatments. Two of those treatments aimed to induce a positive mood (positive treatments), other two aimed to induce a negative mood (negative treatments), and a control treatment served as a reference of participants without induction. The induction was performed by playing a “happy” or “sad” music for participants in the positive or negative treatments respectively. This kind of induction is commonly used in psychology and it has been proven to be as effective as other behavioral mood inductions procedures [Gerrards-Hesse *et al.*, 1994]. Also, during the experiment, two specific games had predetermined results (i.e., win/lose) in order to reinforce the positive/negative mood. A positive/negative induction message was used in one of the two fixed-end games as well (specifically in the first game). In the two positive treatments, as well as in the two negative treatments participants were able to see a graphical history of their results in the previous games. The difference between the two positive treatments and between the two negative treatments was that, in one of both, a ranking of the participants was shown on each game as well as the participant’s position in this ranking.

Figure 6.2 shows the timeline of the experiment, and the configuration of each kind of treatment. On each of the above mentioned treatments, 60

## 6. GENIA<sup>3</sup>'S AGENTS PLAYING BLACK JACK

---

participants randomly assigned to one treatment, played 24 games. As shown in Figure 6.2, all the treatments' first five games were training games. These games were included for the participants to get familiarized with the rules of the game. Each one of these five games was preceded by a question of the individual's earning aspirations for the next game<sup>1</sup>. The subsequent 19 games were not only preceded by a question of aspirations, but also for a question of the participant's perceived autoefficacy (i.e., the extent to which an individual thinks he has the skills to solve a specific task<sup>2</sup> [Bandura, 1996]). Figure 6.2 also shows that, before and after the games 6 and 16, as well as after game 24 (the beginning, the middle, and the end of the set of non-training games), participants needed to fulfill a mood questionnaire. At the beginning of the experiment, participants were also asked to fulfill questionnaires related to their personality and socio-demographic information, and at the end of the experiment a debriefing music was played for those participants under some kind of mood induction, for them to return to a positive mood due to ethical reasons.

At each one of the last 18 games, when either the player or the bank hit a card, the card was randomly selected from the deck. Nevertheless each one of these last 18 games was designed to fulfill some requirements. Figure 6.3(b) shows the three specific kinds of game that were selected for the experiment, and Figure 6.3(a) shows the assignation of the games to one of these three kinds of game. Figure 6.3(b) represents the basic strategy of the Black Jack for a single-deck game where the bank stands when reaching soft<sup>3</sup> 17 (i.e., when reaching 17 points, no matter if it contains an ace or special card whose value can be 1 or 11). This is a strategy that guaranties that, in general, after a big number of games, there is a high probability of ending with more money than at the beginning [Griffin, 1996]. Rows in Figure 6.3(b) represent the total amount of points of the player's cards, and columns represent the points of the bank's visible card. There are some combinations of player's points

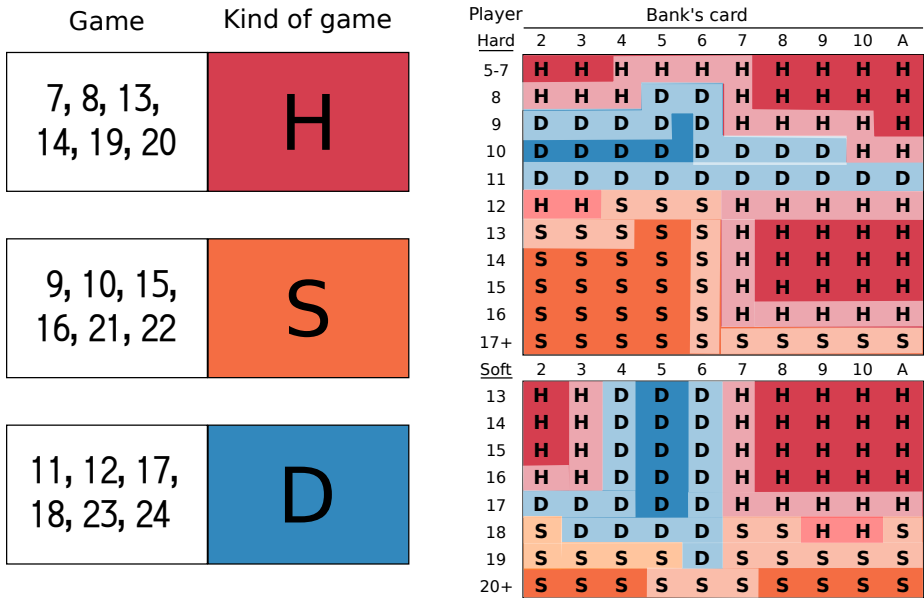
---

<sup>1</sup>For a reference regarding the influence of aspirations on behavior see [Hu *et al.*, 2011].

<sup>2</sup>Albert Bandura called it *efficacy expectations* [Bandura, 1996].

<sup>3</sup>When the player has at least an ace or special card in the game, it is called 'soft' hand. Otherwise it is called 'hard' game.





(a) Assignation of a kind of game to each game. (b) Possible combinations of the bank's and player's cards by kind of game.

Figure 6.3: Classification of games and its assignation to specific kinds of game. H: hit, D: double, S: stand.

## 6. GENIA<sup>3</sup>'S AGENTS PLAYING BLACK JACK

---

and bank's points that are highlighted in soft color. These are the selected combinations that were assigned to the last 18 games<sup>1</sup> as a requirement for the automatic selection of cards (according to Figure 6.3(a)). The selection of these combinations has been done by mainly selecting those combinations in the frontier of two different 'best' plays. In those combinations of cards that are in the frontier it is less clear what is better to do, thus the decision is more complex. The influence of mood on these games should be stronger according to models such as the Affect Infusion Model (AIM)[Forgas, 1995]. The AIM states that mood tends to be more influential when driving responses in situations, and when evaluating situations that are more unanticipated and complex.

### 6.1.3 Measures

Two of the main variables that were measured in the experiment were the participants' personality and mood. In order to measure the participants' personality and mood, it was used the PANAS (Positive and Negative Affect Schedule) questionnaire [Watson *et al.*, 1988] on its Spanish version [Sandín *et al.*, 1999]. This is one of the most used measures in psychology for affective personality and mood assessment [Sandín *et al.*, 1999]. Therefore we used it in the experiment. It is a 20-item measure of positive (*PA*) and negative (*NA*) affect, where each item is self-reported and fairly represents pure descriptor of affect. *PA* and *NA* are independent (and hence almost not correlated) variables. *PA* represents the extent to which an individual feels enthusiastic, inspired, excited, etc., and *NA* represents a dimension including affective states such as distressed, nervous, irritable, etc. [Watson *et al.*, 1988]. These dimensions are known as "descriptively bipolar affectively unipolar" (i.e., a high value on them indicates a high state of the mood, but a low value only represents the relative absence of the corresponding affective implication). These two dimensions can represent either affective states or moods, or personality traits

---

<sup>1</sup>Note that, although game 6 was not a training game, it doesn't have the same requirements of the rest of games that were not training games. In turn, the single requirement that it has is that player's points must be 21 if he is being induced a positive mood. Otherwise bank's points must sum 21 if the player is induced a negative mood.

such as extraversion or neuroticism [Sandín *et al.*, 1999]. Some authors have even found a correspondence between these dimensions and other dimensions of mood in a different scale such as one of the *PAD* (Pleasure, Arousal, Dominance) emotion model [Mehrabian, 1997]. Specifically, we have used the 20-item PANAS questionnaire to measure the participants' affective personality. Also, for simplifying the questionnaires, we have used a reduced version of the PANAS scale to measure the participants mood. This reduced version is the one used in [Van Knippenberg *et al.*, 2010], where it has been proved to fairly capture dispositions to experience negative and positive affective states.

Additionally we were interested on measuring how important it was for an individual to lose or to win a game. Therefore we have used the Schwartz's maximization scale [Schwartz *et al.*, 2002]. It is a 13-item scale which helps to measure a specific trait: an individual's desire to maximize. One may think that as more options one has better. Nevertheless, in [Schwartz *et al.*, 2002] it is shown that there are different scales of satisfaction with the same amount of choices. We have included this scale in order to find out whether this trait helps to predict what specific profits of losses may mean to an individual from an affective point of view. Also this trait helped to predict the participants' decisions during the game.

### 6.1.4 Data Analysis

In this section the analysis of the data collected on the experiment is presented. To this end, several models were adjusted, according to the requirements of the task to be performed. These models helped to predict the mood variation after a game, the bet on a game, and the decisions during a game. The analyzes were conducted using the version 3.3.2 of the R language<sup>1</sup>.

### Predicting Variations on Mood

The first question of Section 6.1.1 is related to the change on an individual's mood after a single game. In Hypothesis 1 we assume that there is a significant change on mood after playing a game. After verifying this hypothesis, we also

---

<sup>1</sup><https://www.r-project.org/>

## 6. GENIA<sup>3</sup>'S AGENTS PLAYING BLACK JACK

---

want to verify Hypotheses 2 by determining the extent to which mood changes depending on the mood induction, and Hypotheses 3 by determining how decisions are taken depending on the participants' mood (positive, negative or neutral).

As shown in Table 6.1, for all kinds of induction there is a significant change on the mean of *PA* and *NA* after the game (won or lost) with respect to the beginning of the game, what confirms Hypothesis 1. These results were obtained by performing paired T-student tests by using games that had similar experimental conditions (i.e., game 16 for participants under a positive and negative induction, and games 16 and 6 for participants without induction<sup>1</sup>. Game 6 of participants under a positive induction (hereinafter positive group) or negative induction (hereinafter negative group), was not used because it contained an additional condition with respect to game 16, which was the induction message. The first four rows of Table 6.1 show the comparison between the *PA* (or *NA*) means after winning (or losing) and the *PA* (or *NA*) means before winning (or losing) for participants without induction (hereinafter control group). The second four rows show the difference between the *PA* means after winning (or losing) and the *PA* means before winning (or losing) of the positive group (which always won the game 16), and the difference between the *NA* means after winning (or losing) and the *NA* mean before winning (or losing) of the negative group (which always lose the game 16). Besides, two similarities on the mean differences can be highlighted: 1) the mean difference of *PA* on the negative group seems similar to the mean difference of *PA* on the control group when they lose, and 2) the mean difference of *NA* on the positive group seems similar to the mean difference of *NA* on the control group when they win. The rest of mean differences doesn't seem similar. An Anova test confirmed similarities on the mean differences of *PA* and *NA* between the control group when winning and the positive group, as well as similarities on the mean differences of *PA* and *NA* between the control group when losing

---

<sup>1</sup>Note that game 16 was controlled in participants with positive or negative induction so that, if the induction was positive, the participant always won game 16, and if the induction was negative, the participant always lose game 16. In participants without induction, both games 6 and 16 were taken into account, because participants won or lost both games randomly.

## 6.1. Black Jack Experiment

Table 6.1: Results of T-test comparisons of the means after and before winning (or losing) by kind of treatment.

Variable	Condition	Game	Mean of the differences	Comparison (P-value)
control group <i>PA</i>	after vs. before winning	16/6	1.029	higher
control group <i>NA</i>	after vs. before winning	16/6	-0.676	lower
control group <i>PA</i>	after vs. before losing	16/6	-0.897	lower
control group <i>NA</i>	after vs. before losing	16/6	0.808	higher
positive group <i>PA</i>	after vs. before winning	16	0.617	higher
positive group <i>NA</i>	after vs. before winning	16	-0.625	lower
negative group <i>PA</i>	after vs. before losing	16	-0.95	lower
negative group <i>NA</i>	after vs. before losing	16	1.342	higher

and the negative group. Thus, with a confidence level of 95%, both *PA* and *NA* changed to the same extent after a won game when participants were listening to a positive music than when not, and both *PA* and *NA* changed to the same extent after a lost game when participants were listening to a negative music than when not. Figure 6.4 shows the mean differences for all considered variables.

In order to confirm Hypotheses 2 and 3, we needed to confirm that groups are indeed in a positive/negative mood compared with the rest. We performed different T-tests that compared the means of the *PA* and *NA* of the positive, negative, and the control groups. Table 6.2 shows these results. The first 6 rows of Table 6.2 compare these three groups before (rows 1 to 3) and after (row 4 to 6) the game where the mood induction was performed (game 6). It can be observed that participants in the positive group were happier than those in the negative group after the 6th game (i.e., they have a higher *PA* after winning), and also happier than the control group. Also, participants in the negative group were sadder than those in the positive group after the 6th game (i.e., they have a higher *NA* after losing), and sadder than the control group. Nevertheless, Table 6.2 shows that when comparing the means of *PA* and *NA* of the positive group with the means of *PA* and *NA* of the control group, they are similar before winning and also similar after winning (results of game 16 for the positive group and of games 6 and 16 for the control group), which supports our previous results of similarities in the extent to which *PA* and *NA* change in these two samples. The mean of *NA* was higher in the

## 6. GENIA<sup>3</sup>'S AGENTS PLAYING BLACK JACK

---

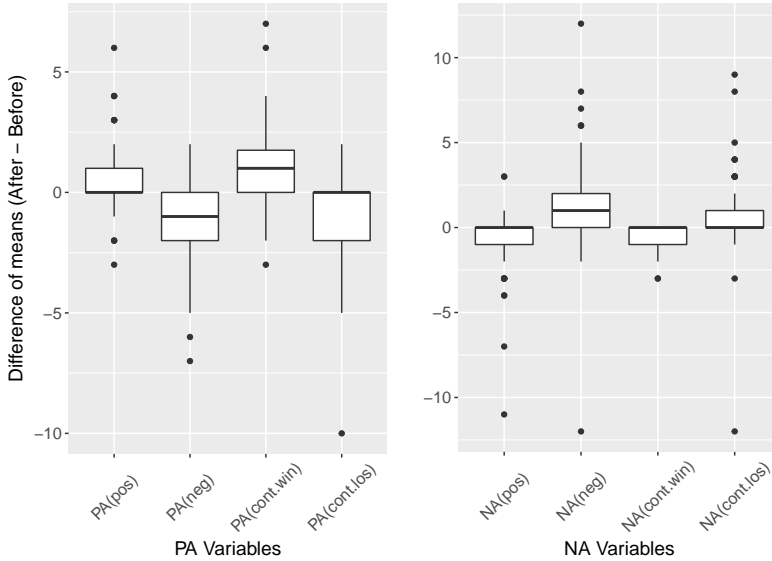


Figure 6.4: Difference of means between after and before a game grouped by *PA* and *NA* (“pos”: positive group, “neg”: negative group, “cont.win”: control group when winning, and “cont.los”: control group when losing).

negative group than the mean of *NA* in the control group before and after losing (results of game 16 for the negative group and of games 6 and 16 for the control group), but they change to the same extent according to the previous results. The mean of *PA* in the negative group was lower than the mean of *PA* in the control group before losing and after losing (and the same applies in relation to the extent that *PA* changes, which is similar in both cases). We can conclude that the music that was used for inducing mood was effective in the first game along with the induction message. Besides, the music may influence the way mood changes after a sequence of games. Nevertheless, music does not have a significant impact on the way the mood changes on a single normal game, confirming the Hypothesis 2. Finally, the last three rows of Table 6.2 confirm that, after the 16th game, participants in the positive group are happier than the rest of participants, and that participants in the negative mood are sadder than the rest of participants.

After examining the Spearman correlation of several variables that were part

## 6.1. Black Jack Experiment

Table 6.2: Comparison of the means of 11 samples of the positive, negative, and control groups with a confidence level of 95%.

Samples	<i>P</i> As	<i>N</i> As
Pos (*6) vs. Neg (*6)	= Means: 9.94 vs. 9.50	= Means: 3.75 vs. 4.08
Pos (*6) vs Cont (*6)	= Means: 9.94 vs. 9.48	= Means: 3.75 vs. 3.68
Neg (*6) vs Cont (*6)	= Means: 9.50 vs. 9.48	= Means: 4.08 vs. 3.68
Pos (6*) vs Neg (6*)	> Means: 10.38 vs. 8.70	< Means: 3.50 vs. 4.83
Pos (6*) vs Cont (6*)	> Means: 10.37 vs. 9.13	< Means: 3.50 vs. 4.08
Cont (6*) vs Neg (6*)	= Means: 9.13 vs. 8.70	< Means: 4.08 vs. 4.83
Pos (*16) vs Cont.win (*16/6)	= Means: 8.93 vs. 8.97	= Means: 4.75 vs. 4.26
Pos (16*) vs Cont.win (16/6*)	= Means: 9.55 vs. 10.00	= Means: 4.13 vs. 3.59
Neg (*16) vs Cont.los (*16/6)	< Means: 7.74 vs. 9.12	> Means: 6.03 vs. 4.23
Neg (16*) vs Cont.los (16/6*)	< Means: 6.79 vs. 8.22	> Means: 7.38 vs. 5.04
Pos (16*) vs Neg (16*)	> Means: 9.55 vs. 6.79	< Means: 4.13 vs. 7.38
Pos (16*) vs Cont (16*)	> Means: 9.55 vs. 8.40	< Means: 4.13 vs. 5.10
Cont (16*) vs Neg (16*)	> Means: 8.40 vs. 6.79	< Means: 5.10 vs. 7.38

Samples are represented according to the following notation:  $\langle group \rangle ([*] \langle games \rangle [*])$  where *group*, can be one of: Pos (positive group), Neg (negative group), Cont (control group), Cont.win (winning games of the control group), and Cont.los (losing games of the control group); *game* can be 6, 16 or both (game (s) of the samples); and \* can be situated before/after the game number (indicating that these are values before/after the game or games of analysis).

## 6. GENIA<sup>3</sup>'S AGENTS PLAYING BLACK JACK

Table 6.3: Variables with better correlation and best model fitted for  $vPA$  and  $vNA$ .

Dependent variable	Independent variables	Correlation coefficients	Best model <sup>a</sup>	adj. $R^2$
Variation of $PA$	E	0.43	$-1.87 + 3.78E * P_{NA}^c$	0.10
	E * MS	0.42		
	$P_{PA}^b * E$	0.42		
	$P_{NA}^c * E$	0.42		
	$P_{PA}^b * MS * E$	0.42		
	$P_{NA}^c * MS * E$	0.42		
	$MS_2 * E$	0.41		
	$MS_4 * E$	0.43		
Variation of $NA$	E	-0.46	$1.77 - 4.20E * MS + 0.75P_{PA}$	0.10
	E * MS	-0.45		
	$P_{PA}^b * E$	-0.45		
	$P_{NA}^c * E$	-0.46		
	$P_{PA}^b * MS * E$	-0.44		
	$P_{NA}^c * MS * E$	-0.45		
	$MS_2 * E$	-0.43		
	$MS_4 * E$	-0.43		

E: Earnings, MS: Maximization scale,  $P_{PA}$ : Personality's PA,  $P_{NA}$ : Personality's NA,  $MS_2$ : 2<sup>nd</sup> component of the MS, and  $MS_4$ : 4<sup>th</sup> component of the MS.

<sup>a</sup>According to lowest MSE in 10 iterations and using different combinations of training and test sets.

<sup>b</sup>Only the components 'Decidido', 'Activo', and 'Dispuesto'.

<sup>c</sup>Only the components 'Culpable', 'Avergonzado', and 'Asustado'.



of the game and the variation of the mood variables  $PA$  and  $NA$  (hereinafter  $vPA$  and  $vNA$ ), we determined the highest correlation coefficients ‘ $r$ ’. They were moderate correlations<sup>1</sup> between  $vPA$  or  $vNA$  and the features that involved earnings in the game, the participant’s maximization scale, and some other personality traits. The best correlation with  $vPA$  ( $r = 0.43$ ) was with earnings in the game, and also with the multiplication of earnings with other features. The same applied to the best correlation with  $vNA$  ( $r = -0.46$ ) which was also with earnings and with the feature resulting from the multiplication of earnings with the personality’s negative affect. Other variables also had similar correlations with  $vPA$  and  $vNA$ . Table 6.3 shows the variables and the corresponding correlation coefficients. With this information we have fitted two linear regression models in order to predict  $vPA$  and  $vNA$ . The column ‘Best model’ contains the best models for predicting  $vPA$  and  $vNA$ . Adjusted R squares (adj.  $R^2$ ) were 0.103 ( $vPA$  model) and 0.101 ( $vNA$  model) for 419 training samples and 181 test samples. We have selected these models from a set of models (built with different combinations of scaled features on the training and test sets) according to the lowest MSE.

### Grouping Samples by Positive, Negative and Neutral Mood

Previous results confirmed that, after playing a game, participants that won the game were happier than the rest of participants, and that participants that lost the game were sadder than the rest of participants. Considering this, we made three groups of participants: those with a positive mood, those with a negative mood, and those with a neutral mood. This grouping allowed to verify Hypothesis 3 and hence to determine whether mood influenced decisions and how. In order to classify each individual’s mood a mechanism was required. To this end, we fitted a decision tree with the variables  $PA$  and  $NA$  with tree possible classes: positive mood (PosMood), negative mood (NegMood), and neutral mood (NeutMood). Figure 6.5 shows a graphical representation of the decision tree.

---

<sup>1</sup>Following the Cohen’s convention to interpret coefficients [Cohen, 1988], a small correlation has a coefficient above 0.10, a moderate correlation has a coefficient above 0.30, and a large correlation has a coefficient above 0.50.

## 6. GENIA<sup>3</sup>'S AGENTS PLAYING BLACK JACK

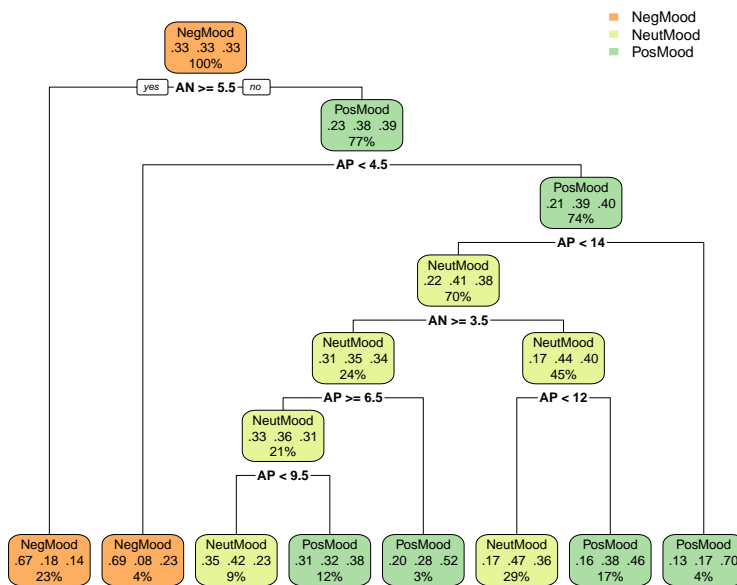


Figure 6.5: Decision tree for classifying each individual's mood in positive (PosMood), negative (NegMood), or neutral (NeutMood).

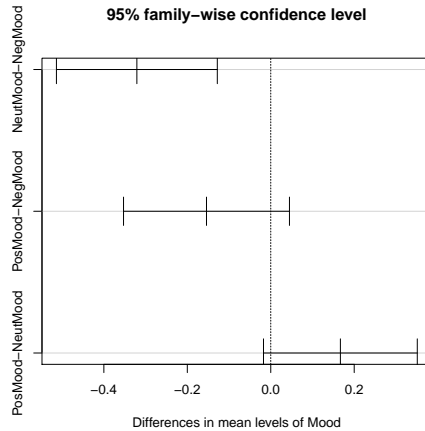


Figure 6.6: Mean difference and confidence intervals with a confidence level of 95% in a pairwise comparison of bet by kind of mood.

720 observations were used, where 240 were considered in a positive mood (mood of 120 individuals after winning games 6 and 16 in the positive induction), 240 in a negative mood (mood of 120 individuals after losing games 6 and 16 in the negative induction), and 240 in a neutral mood (mood of the previous 240 individuals at the beginning of the experiment). These observations were considered to be in a positive, negative, or neutral mood on the basis of the results of Table 6.2. This decision tree (hereinafter mood DT) was used to determine whether an individual was in a positive, negative, or neutral mood.

### Predicting Bet

In order to predict participants' bets, we included games 6, 7, 16, and 17 (1200 observations). The data of the mood before the game for these four games was available, thus allowing to determine whether and how mood influenced the corresponding bets. We assigned each of the 1200 observations to one of the PosMood, NegMood, or NeutMood classes according to the mood DT. In order to determine whether the mean differences between the bets of the PosMood, NegMood, or NeutMood observations were significant, an ANOVA test was performed. The test didn't show significant differences in the means

## 6. GENIA<sup>3</sup>'S AGENTS PLAYING BLACK JACK

---

between PosMood and NegMood observations (P-value of 0.16), or between PosMood and NeutMood observations (P-value of 0.08). Nevertheless there were significant differences between the bets of NegMood and NeutMood observations (P-value of 0.00). This means that bets of participants in a negative mood are different from bets of participants in a neutral mood. Figure 6.6 shows these mean differences and confidence intervals of a pairwise comparison of groups. It shows that the difference of NeutMood and NegMood bets' means is significantly far from the zero value. Besides, this mean difference is negative, meaning that participants in a negative mood tend to have higher bets than participants in a neutral mood. It is also observable that, although not statistically significant, participants in a negative mood also tend to have higher bets than participants in a positive mood, and that participants in a positive mood tend to have higher bets than participants in a neutral mood. This supports the Hypotheses 3 presented in Section 6.1.1.

Considering that the bets of participants with different mood can be different, we have grouped observations by kind of mood in order to fit a linear regression model by kind of mood for predicting the bet. In order to determine the variables with more influence on the bet, we have built a correlation matrix, considering games 6, 7, 16, and 17. In general, the variables that showed the highest correlations for the samples of games 7, 16, and 17 included 'bet of previous game' ( $r = 0.72$ ), and 'balance'<sup>1</sup> ( $r = -0.19$ ). Another variable that also showed a strong correlation was 'aspirations' ( $r = 0.62$ ), but it was not included in the models because it anticipated the participant bet intentions before betting. By analyzing correlations by kind of mood we obtained the results shown on Table 6.4. These results are divided into two samples: those of games 7, 16, and 17, and those of the game 6, because not all variables were part of game 6 (e.g., 'earnings in previous game' or 'bet of previous game'). A training and a test set, with scaled values, were used to fit the models of games 7, 16, and 17 (177/119 for observations in a positive mood, 181/121 for observations in a neutral mood, and 181/121 for observations in a negative mood), and a single set was used for training and testing the model of game 6

---

<sup>1</sup>Amount of money accumulated in the wallet.

## 6.1. Black Jack Experiment

Table 6.4: Variables with better correlation and best models fitted for ‘bet of the game’.

Dependent var.	Independent var.	Correlation coefficients				Best models <sup>a</sup>	MSE
		Pos Mood	Neut Mood	Neg Mood	All		
Bet in game 6	$P_{PA}$	0.19	0.05	0.14	0.11		
	$P_{PA}^b$	0.18	0.02	0.27	0.11	<b>All:</b>	
	$P_{NA}$	-0.16	-0.04	-0.06	-0.07	$1.74 + 1.06P_{PA}^b - 0.46P_{NA} +$	1.028
	$PA$	0.21	0.07	0.24	0.08	$1.11NA$	
	$NA$	-0.29	-0.03	0.25	-0.01		
Bet in games 7, 16, and 17	$PB$	0.69	0.73	0.74	0.72	<b>PosMood:</b> $1.20 + 2.48PB - 0.26NA$	0.459
	$CB$	-0.06	-0.08	-0.34	-0.19	<b>NeutMood obs:</b> $1.67 + 2.5PB - 1.37CB$	0.371
	$PA$	0.08	-0.03	0.01	-0.01	<b>NegMood obs:</b> $1.56 + 2.65PB - 1.10CB$	0.431
						<b>All:</b>	
	$NA$	-0.11	0.08	0.08	0.06	$1.31 + 2.74PB - 0.91CB - 0.21NA$	0.509

$P_{PA}$ : Personality’s PA,  $P_{NA}$ : Personality’s NA, MS: Maximization scale, PB: Bet in previous game, CB: Current balance

<sup>a</sup>According to the MSE in 10 different combinations of training and test sets.

<sup>b</sup>Only the components ‘Energetico’ and ‘Orgulloso’.

<sup>c</sup>Only the components ‘Decidido’, ‘Entusiasmado’, and ‘Dispuesto’.

(104 observations in a positive mood, 158 for observations in a neutral mood, and 38 for observations in a negative mood).

As shown in Table 6.4, for games 7, 16, and 17, the most influential variable is ‘bet of previous game’, suggesting that participants had a tendency to bet in a similar way in two consecutive games. Besides, other small and moderate correlations can be observed with personality’s  $NA$  in the PosMood sample, and with ‘current balance’ in the NegMood sample. The variable that showed the best correlations with ‘bet of the game’ for game 6 was personality’s  $PA$ , which was a small correlation. After fitting several linear regression models with the best correlated variables, we selected those with the best MSE that improved the MSE of predicting the mean. The best models for each kind of mood are also shown in Table 6.4. The amount of observations for game 6 didn’t allow to build a reliable predicting model by kind of mood, thus the best model for the whole sample was fitted. For observations of games 7, 16, and 17 three models were fitted, one for each of the PosMood, NeutMood, and NegMood classes.

## 6. GENIA<sup>3</sup>'S AGENTS PLAYING BLACK JACK

Table 6.5: Best predicting models of whether to hit, to stand or double by kind of mood and by kind of game.

Sample		Best models <sup>a</sup>	Test <i>ce</i>	Sample size (training/test)
Whole sample		VPC + SH	0.408	528/353
PosMood		VPC + SH + VBC + CB + PE	0.365	172/115
NegMood		VPC + SH + VBC + CB + PA + NA	0.356	177/118
NeutMood		VPC + SH	0.283	179/120
The best is to HIT	All	VPC + HI	0.409	170/115
	PosMood	VPC + VBC + NA	0.518	110 <sup>b</sup>
	NegMood	VPC + VBC + PE + PA + NA	0.507	69 <sup>b</sup>
	NeutMood	VPC + SH + VBC + B	0.406	106 <sup>b</sup>
The best is to STAND	All	VPC + SH	0.287	177/119
	PosMood	VPC + B	0.321	84 <sup>b</sup>
	NegMood	VPC + PE	0.363	102 <sup>b</sup>
	NeutMood	VPC + B + PA + NA	0.309	110 <sup>b</sup>
The best is to DOUBLE	All	VPC + SH	0.352	178/122
	PosMood	VPC + SH + B + PA + NA	0.301	93 <sup>b</sup>
	NegMood	VPC + SH	0.411	124 <sup>b</sup>
	NeutMood	VPC + SH + VBC + NA	0.349	83 <sup>b</sup>

VPC: total value of the player's cards, VBC: total value of the Bank's cards, CB: current balance, SH: Soft hand, HI: Historical information of the previous games shown in the screen, PE: previous earnings, B: current bet.

<sup>a</sup> Those with the smallest Test *ce* and using balanced categories.

<sup>b</sup> The whole sample is used for training and for testing.

### Predicting First Decision

According to the rules of our version of Black Jack, when a game starts, the player needs to bet before the cards are shown. Then, two cards are given to the player and two cards to the bank. After the player sees the cards<sup>1</sup>, the player can decide one of three options: to double the bet, to hit a card, or to stand. Thus, after betting, this is the first decision that the player must perform.

In order to predict this first decision a multinomial regression model was fitted, which best predicted this decision. In general, the models that best predicted the first decision in the game included, among others, the following

<sup>1</sup>He can see his two cards and only one of the bank's cards because the other card is hidden

## 6.1. Black Jack Experiment

Table 6.6: Best predicting models of whether to hit or to stand by kind of mood and by kind of game.

Sample	Best models <sup>a</sup>	Test <i>ce</i>	Sample size (training/test)
Whole sample	VPC + SH + VBC + NA	0.080	204/138
PosMood	VPC + NA	0.111	117 <sup>b</sup>
NegMood	VPC + VBC	0.134	97 <sup>b</sup>
NeutMood	VPC + VBC + PA + NA	0.109	128 <sup>b</sup>
The best is to HIT	VPC + B + PA + NA	0.152	112 <sup>b</sup>
The best is to STAND	VPC + CB + NA	0.065	137/93

VPC: total value of the player's cards, SH: Soft hand, VBC: total value of the Bank's cards, B: bet of the game, CB: current balance.

<sup>a</sup> Those with the smallest Test *ce*.

<sup>b</sup> The whole sample is used for training and for testing.

variables: user's and bank's cards, whether the user's cards included an ACE ('soft hand'), 'current balance', and the mood before starting the game. We used the data of games 7, 16, and 17 (in total 881 observations), because they contained information about the user's first play, about the mood before the game, and they belonged each to a kind of game (see Figure 6.3(b)). As described in Section 6.1.2, the experiment was properly designed so that games 7, 16, and 17 responded to situations where the best rational option was to hit, to stand, and to double respectively. Table 6.5 shows the results for observations grouped by kind of mood (PosMood, NegMood, and NeutMood), by kind of game, or by both. The selection of the best predicting models shown in Table 6.5 was done according to the minimum test classification error *ce*. In general the best models were those where observations were grouped by kind of mood, and grouping observations by kind of situation or by both kind of situation and kind of mood didn't improve the results.

### Predicting When to Stand

Once the player has made the first decision, if he didn't double or stand, he must decide again whether to hit a new card or to stand. For predicting these subsequent decisions, we performed the same steps as for predicting the first

## 6. GENIA<sup>3</sup>'S AGENTS PLAYING BLACK JACK

---

decision, but using a binomial logistic regression in this case. The results are shown in Table 6.6, where observations were grouped by kind of mood and by kind of game<sup>1</sup>. In general, and considering the size of the samples, the *ce* of the models that resulted of grouping by kind of mood or by kind of game, didn't improve the *ce* of the model fitted for the whole sample.

### 6.2 Agents Playing Black Jack

In this section we use the results of the Black Jack experiment in order to illustrate how to build three kind of agents: rational agents, agents compliant with the default design of *GenIA*<sup>3</sup>, and agents compliant with an extended version of the *GenIA*<sup>3</sup>'s default design. In the extended version we describe how the default design can be extended for including more detailed information of the results of the experiment. Three simulations were performed, one by kind of agent. All simulations had the same characteristics of the experiment (i.e., each one included 300 agents, five kinds of treatment, agents played 19 games, and the characteristics of the games for each treatment were the same as those of the experiment). The results obtained for the three kind of agents are compared regarding their decisions on the games, and regarding the mood fluctuations when possible.

#### 6.2.1 Rational Agents

The best models fitted in Section 6.1.4 included, to a large extent, mood or personality related variables, with the exception of the models for predicting the first play on the game. The models fitted for predicting the first play on the game included the value of the player's cards and whether the game was a 'soft hand'. These variables, and also the value of the bank's cards, are precisely the variables used in the basic strategy (see Figure 6.3) to find out which is the best play to do for a given combination of player's and bank's cards. Considering that, in an environment where no human factors can have

---

<sup>1</sup>The kind of game was determined again, now considering the card the player previously obtained.



```

1  @pstand
2  +!play(VPC,VBC,As)
3      :    currentGame(CG) &
4          currentPlay(CP) &
5          standPlays(SP) &
6          .nth(I,SP,p(VPC,VBC,As,_)) & I>=0 &
7          croupier(S)
8  <- .print("STAND with ",VPC," and bank has ",VBC,
9         " on play ",CP," of game ",CG);
10     if(CP == 1 & VPC == 21){
11         !endOfGame(VPC,VBC);
12     }else{
13         .send(S,achieve,playerStands(VPC));
14     }.

```

Figure 6.7: Code of the plan for standing of the rational agents.

influence over decisions, the basic strategy is the best performance<sup>1</sup>, thus our rational agents were designed for them to play the basic strategy. Besides, considering that there weren't possible losses, regardless what the bet was, rational agents always bet the maximum possible amount of money (which was €4.0).

Figure 6.7 shows an extract of the code of the rational agents. Specifically, Figure 6.7 shows the plan where the agent executes the actions for standing. The context of the plan ensures that the values of the player's cards (VPC), the value of the bank's cards (VBC) and whether the current game (CG) is a soft hand (As) correspond to a play where the best is to stand according to the basic strategy (lines 5-6). If the agent didn't reach 21 points on the first play, it sends a message to the bank for it to start its turn to play (meaning that the agent stands).

## 6. GENIA<sup>3</sup>'S AGENTS PLAYING BLACK JACK

Table 6.7: Alternative model for predicting ‘bet of the game’ in the first non-training game without personality variables.

Dependent variable	Independent variables	Correlation coefficients				All	Best model <sup>a</sup>	MSE
		Pos Mood	Neut Mood	Neg Mood				
Bet in game 6	<i>PA</i> <i>NA</i>	0.21 -0.29	0.07 -0.03	0.24 0.25	0.08 -0.01	<b>All:</b> $1.56 + 0.95PA + 0.88NA$	1.046	

PA: Mood positive affect, NA: Mood

<sup>a</sup>According to the MSE in 10 different combinations of training and test sets.

### 6.2.2 Agents Following the GenIA<sup>3</sup>'s Default Design

Agents following the default design of *GenIA*<sup>3</sup>, were specified by using the classes it provides (see Figure 4.5). Specifically the class for managing the processes of the agent’s affective cycle, contains default mechanisms for appraisal, or affective state updating, among others, which were used without modifications. As described in Section 4.5, the default design of *GenIA*<sup>3</sup> works with the PAD measure of affect Mehrabian [1996b]. In order to find a correspondence between the affective state measured in the PANAS scale, and the affective state used in the default design of *GenIA*<sup>3</sup> expressed in the PAD scale, the regression models proposed in [Mehrabian, 1997, p. 342 and 344] were used<sup>1</sup>. These correspondences allowed to initialize the agents’ affective state with the initial affective state of the participants of the experiment, and also to obtain the PANAS variables (*PA* and *NA*) on the basis of the agent current affective state expressed in the PAD scale. Regarding the agent’s personality, in *GenIA*<sup>3</sup>’s default design it is represented in the Five OCEAN scale [McCrae & John, 1992]. Considering that, to the best of our knowledge, there isn’t still any report of the correspondence between the PANAS scale of affective personality and the OCEAN scale in both directions<sup>2</sup>, personality

<sup>1</sup>The basic strategy has been obtained from probabilistic studies for a single deck game, and with the characteristics of the games of the experiment [Griffin, 1996, 18], and it is shown in Figure 6.3(b)).

<sup>1</sup>Specifically the regression equations for deriving *PA* and *NA* values in the PANAS scale with *P*, *A*, and *D* as independent variables, and the regression equations for deriving *P* and *A*, and *D* values in the PAD scale with *PA*, and *NA* as independent variables, for all the participants of the study.

<sup>2</sup>Some related studies have been performed such as [Zanon *et al.*, 2013]. Nevertheless, it is hard to find a general and bidirectional correspondence between the OCEAN and the

```

1  MAS defaultAgents {
2    infrastructure: Centralised
3    agents:
4      bjbank;
5      bjagent
6        beliefBaseClass customClasses.BJdemoBB
7          ("com.mysql.jdbc.Driver",
8            "jdbc:mysql://DBserver.dsic.upv.es/BJ-DB",
9            "BJ-user", "BJ-password",
10           "[subjects(12), mood(9)]")
11       agentClass jason.Em.asSemantics.EmAgent
12       agentArchClass jason.Em.architecture.EmAgArch
13       emEngineClass jason.Em.architecture.DefaultEmotionEngine
14       personalityClass jason.Em.architecture.OCEANPersonality
15       affStateClass jason.Em.architecture.PAD          #300;
16
17       affcategories:
18         neg1( <-1.0:0.25> , < -1.0: 1.0 > , <-1.0:1.0 > ),
19         neg2( <0.25:0.53> , <-0.32: 1.0 > , <-1.0:1.0 > ),
20         pos1( <0.25:0.53> , < -1.0:-0.63> , <-1.0:1.0 > ),
21         pos2( <0.57:0.9 > , < -1.0: 1.0 > , <-1.0:0.47> ),
22         pos3( <0.53:0.9 > , < -1.0: 1.0 > , <0.64:1.0 > ),
23         pos4( < 0.9:1.0 > , < -1.0: 1.0 > , <-1.0:1.0 > ),
24         neu1( <0.25:0.53> , <-0.63:-0.32> , <-1.0:1.0 > ),
25         neu2( <0.53:0.9 > , < -1.0: 1.0 > , <0.47:0.64> ),
26         neu3( <0.53:0.57> , < -1.0: 1.0 > , <-1.0:0.47> );
27 }

```

Figure 6.8: MAS project configuration for agents following the *GenIA*<sup>3</sup>'s default design.

was not considered in the agents following the default design. Due to this, it was used an alternative model for predicting the bet of the game, that didn't include personality-related variables, which is shown on Table 6.7.

With the exception of the decision of the bet to perform in the first non-training game (game 6), the decisions of agents following either the default design or the extended design are equally determined on the basis of the models of Section 6.1.4. Figures 6.8 and 6.9 show an extract of the code of the MAS project configuration and of the agent following the *GenIA*<sup>3</sup>'s default design

---

PANAS for affective personality scales, because it generally depends on factors such as age and cultural differences.

## 6. GENIA<sup>3</sup>'S AGENTS PLAYING BLACK JACK

---

```
1 concerns__(V) :- (earnings(E,CG) & (E<=4 & E>=-4) & V=(E + 4.0)/8.0) |
2                   (earnings(E,CG) & (E> 4 | E< -4) & V=(E + 8.0)/16.0) .
3
4 @pdec_neg_aff_play1[affect__(neg1,neg2)]
5 +!play(VPC,VBC,As)
6   :      currentGame(CG) &
7         currentPlay(1) &
8         croupier(S) &
9         currentBet(CB)
10  <-    .get_affst ([P,A,D]);
11        customClasses.getPANASfromPAD(P,A,D,[PA,NA]);
12        CoefDouble = [9.80 , -0.71, 1.02, 0.02 , 0.15 , -0.02, 0.07];
13        CoefHit     = [10.76, -0.77, 0.70, -0.04, -0.11, 0.11 , 0.10];
14        customClasses.getPlay
15          ([VPC,As,VBC,CB,PA,NA],CoefDouble,CoefHit,Play);
16        !applyPlay(Play,VPC,VBC,As) .
```

Figure 6.9: An extract of the code of the agent following the *GenIA*<sup>3</sup>'s default design.

respectively. Figure 6.8 shows that the MAS includes 301 agents, the 'bjbank' (representing the bank), and 300 'bjagent' (see line 15), which are the agents following the default design representing each participant of the experiment. Lines 6 to 10 show the configuration for the agent to connect to the database and to load its initial mood and its kind of treatment. Figure 6.8 also shows the MAS affective categories (see lines 17 to 26). These affective categories are expressed in the PAD scale and were determined by mapping the mood in the PANAS scale from the data to the PAD scale, and by training a decision tree following the same steps of Section 6.1.4.

On the other hand, Figure 6.9 shows an extract of the code of the agent following the *GenIA*<sup>3</sup>'s default design. The agent's concerns are shown in lines 1-2. They contain the single variable *E* (earnings in the last game). This variable showed a relevant positive correlation with the increment of the positive affect and a relevant negative correlation with the increment of the negative affect. Thus it offers clues of the desirability of the situation when the game ends, which is, in fact, what concerns express. Lines 7-15 of Figure 6.9 contain the plan for making the first play when it has a negative mood. The

condition for a negative mood is expressed in line 4 as a plan annotation. The internal action<sup>1</sup> `.get_affst([P,A,D])` of line 10 is included in the *GenIA*<sup>3</sup> platform together with the internal action `.set_affst([P,A,D])`. They help to recover and to set the current affective state respectively. Besides, the internal actions `getPANASfromPAD(...)` and `getPADfromPANAS(...)` were included in the simulation for converting the PAD values of the affective state to a PANAS scale and vice versa. Also the internal action `getPlay(...)` returns, in the variable `Play`, the next play to be performed ('double', 'hit', or 'stand'), given a list of values of the variables of the model, the coefficients of 'double', and the coefficients of 'hit'. Finally, the agent performs the play on the variable `Play` through the goal `!applyPlay(...)` of line 15.

### 6.2.3 Agents Following the *GenIA*<sup>3</sup>'s Extended Design

Agents following the extended design weren't defined with *GenIA*<sup>3</sup>'s default classes, but specific classes were created (which implemented *GenIA*<sup>3</sup>'s interfaces). Figure 6.10 shows the MAS project configuration for the agents following the *GenIA*<sup>3</sup> extended design. Specifically, the class `bjEmEngine` managed the emotion engine, `bjPersonality` managed the agent's personality, and `bjPANAS` managed the affective state. In `bjEmEngine`'s appraisal and affect updating processes, the models fitted for determining mood variations were included. On the other hand, `bjPersonality` included those personality traits that are used by the models for predicting mood variations and decisions of the game, and `bjPANAS` included the necessary procedures for managing the affective state in the PANAS scale. The ranges for the variables were the same used in the experiment. Figure 6.11 shows an extract of the code of the agent following the extended design similar to that of Figure 6.9. It can be seen that concerns also include the variable *E* (earnings in the last game). In this example it was not used a scaled value for concerns because this, and other required transformations, are done in the emotion engine's customized class, which also considers the agent's personality. The agent's personality is

---

<sup>1</sup>Internal actions allow to perform processing by accessing to legacy code, which is Java in this case. For more information see [Bordini *et al.*, 2007].

## 6. GENIA<sup>3</sup>'S AGENTS PLAYING BLACK JACK

---

```
1  MAS extendedAgents {
2    infrastructure: Centralised
3    agents:
4      bjbank;
5      bjagent
6      beliefBaseClass customClasses.BJdemoBB
7        ("com.mysql.jdbc.Driver",
8         "jdbc:mysql://DBserver.dsic.upv.es/BJ-DB",
9         "BJ-user", "BJ-password",
10        "[subjects(12), mood(9), personality(6)]")
11     agentClass jason.Em.asSemantics.EmAgent
12     agentArchClass jason.Em.architecture.EmAgArch
13     emEngineClass jason.Em.architecture.bjEmEngine
14     personalityClass jason.Em.architecture.bjPersonality
15     affStateClass jason.Em.architecture.bjPANAS      #300;
16
17     affcategories:
18       neg1( < 3.0:15.0> , <5.5:15.0> ),
19       neg2( < 3.0: 4.5> , <3.0:5.5 > ),
20       pos1( < 9.5:14.0> , <3.5:5.5 > ),
21       pos2( < 4.5: 6.5> , <3.5:5.5 > ),
22       pos3( <12.0:15.0> , <3.0:5.5 > ),
23       pos4( <14.0:15.0> , <3.0:5.5 > ),
24       neu1( < 6.5: 9.5> , <3.5:5.5 > ),
25       neu2( < 4.5:12.0> , <3.0:3.5 > ),
26       neu3( < 3.0:15.0> , <3.0:15.0> );
27 }
```

Figure 6.10: MAS project configuration for agents following the *GenIA*<sup>3</sup>'s extended design.

```

1 concerns__(V) :- earnings(E,CG) & V=E.
2
3 @pdec_neg_aff_play1[affect__(neg1,neg2)]
4 +!play(VPC,VBC,As)
5   :   currentGame(CG) &
6       currentPlay(1) &
7       croupier(S) &
8       currentBet(CB)
9   <- .get_affst([PA,NA]);
10  CoefDouble = [9.80 , -0.71, 1.02, 0.02 , 0.15 , -0.02, 0.07];
11  CoefHit    = [10.76, -0.77, 0.70, -0.04, -0.11, 0.11 , 0.10];
12  customClasses.getPlay
13      ([VPC,As,VBC,CB,PA,NA],CoefDouble,CoefHit,Play);
      !applyPlay(Play,VPC,VBC,As) .

```

Figure 6.11: An extract of the code of the agent following the *GenIA*<sup>3</sup>'s extended design.

read from the data base in the initialization process of the customized emotion engine. As shown in Figure 6.11, the agent's first play is determined in the same way as for the agent following the default design. There is a single difference, which is that the affective state isn't transformed, neither during the mood initialization, or inside of the agent's plans since it is no longer necessary because the affective state is expressed in the same scale used in the experiment.

## 6.3 General Results

The results of each simulation will be presented by mainly keeping in mind two issues: the extent to which the agents were able to experience a mood similar to that of the participants of the experiment, and the extent to which the agents were able to perform similar decisions. Sections 6.3.1 and 6.3.2 describe the results for each case.

## 6. GENIA<sup>3</sup>'S AGENTS PLAYING BLACK JACK

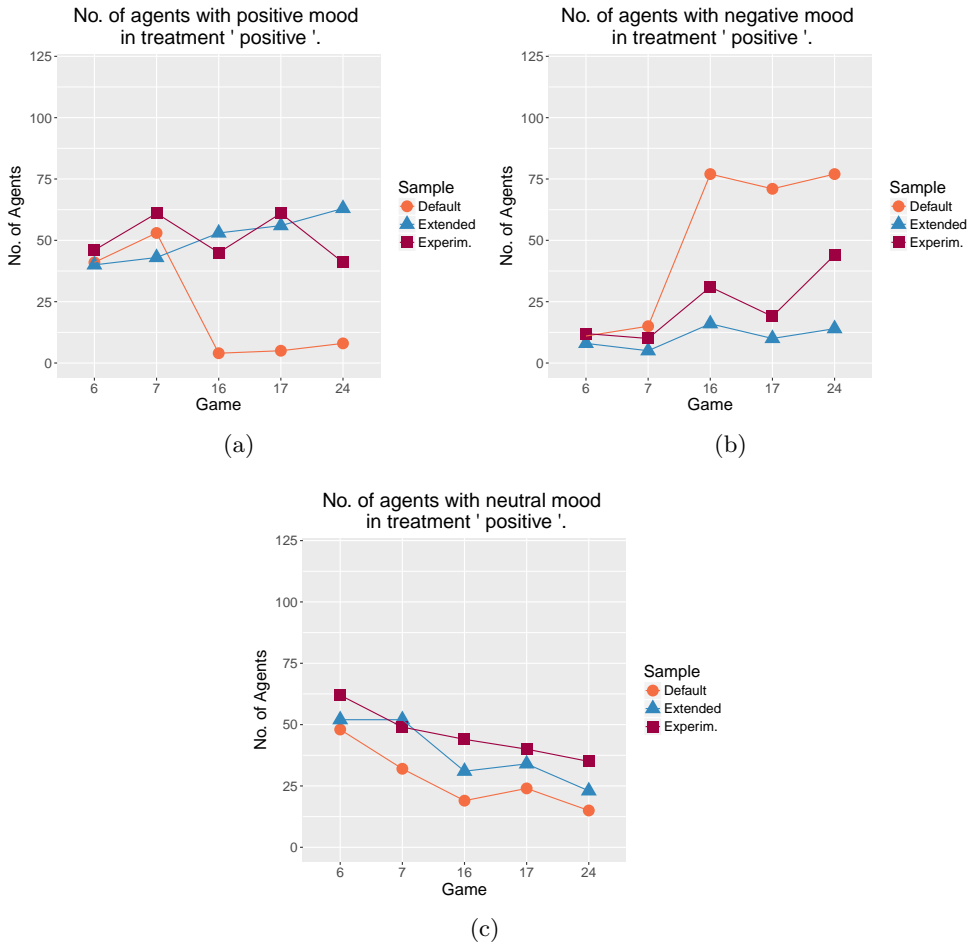


Figure 6.12: Comparison of the number of agents with positive (a), negative (b), and neutral mood (c) in games 6, 7, 16, 17 and 24 in the positive treatment.



### 6.3.1 Agent's Mood Experience

Figures 6.12 and 6.13 show, for each kind of simulation (rational agents, agents following the default design, and agents following the extended design), the number of agents with positive, negative, and neutral mood when playing the games 6, 7, 16, 16, and 24. These are the non-training games on which the participant's mood information was requested in the experiment. Figure 6.12 shows these results for the positive mood induction (or treatment), and Figure 6.13 shows these result for the negative treatment. Agents (and participants) in the positive treatment always won games 6 and 16. Similarly, agents (and participants) in the negative treatment always lost games 6 and 16. Figures 6.12(a) and 6.13(b) show that, in those games were the agents won (games 6 and 16), the number of agents with positive mood generally tends to increase, and in those games were they lost (games 6 and 16), the number of agents with negative mood generally tends to increase. This was a result obtained in the experiment that both simulations have properly represented. In Figures 6.12(c) and 6.13(c), the data of the experiment shows that there is a tendency for the number of participants in a neutral mood, to decrease throughout the experiment, meaning that they tended to get either a positive or a negative mood on each game. This has also been a result of the simulations with the agents.

Besides, Figures 6.12 and 6.13 show that better approximations of the number of agents with a specific mood are performed by the extended simulation in the positive treatment (Figure 6.12) that in the negative treatment (Figure 6.13). In the negative treatment, the result of the default simulation were closer to the results of the experiment than the extended simulation. This means that, the default simulation performs better when predicting negative affective states than when predicting positive affective states. A possible explanation for this is that, in the default simulation the proportion of lost games and won games (725 lost and 483 won) is higher than this proportion in the extended simulation (685 lost and 519 won) and higher than in the experiment (732 lost and 702 won). Besides, in the *GenIA*<sup>3</sup>'s default design, mood is updated according to [Gebhard, 2005], where the *pull and push* function is used. This

## 6. GENIA<sup>3</sup>'S AGENTS PLAYING BLACK JACK

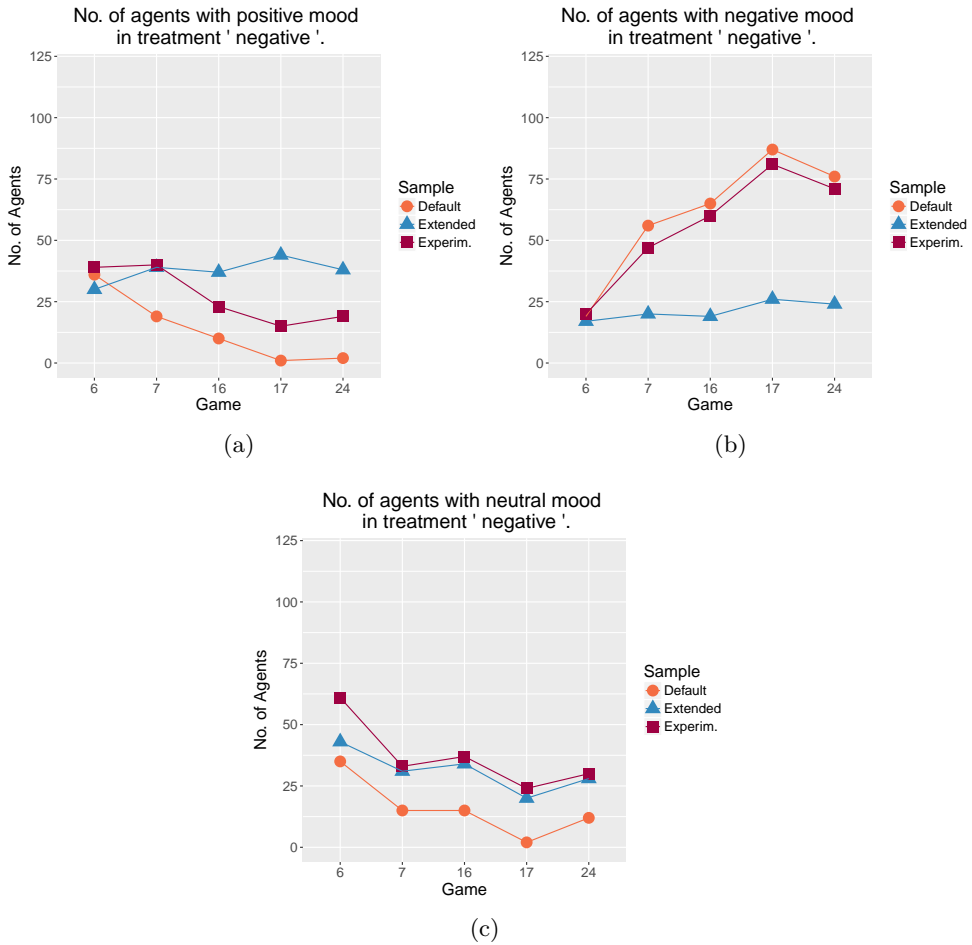


Figure 6.13: Comparison of the number of agents with positive (a), negative (b), and neutral mood (c) in games 6, 7, 16, 17 and 24 in the negative treatment.

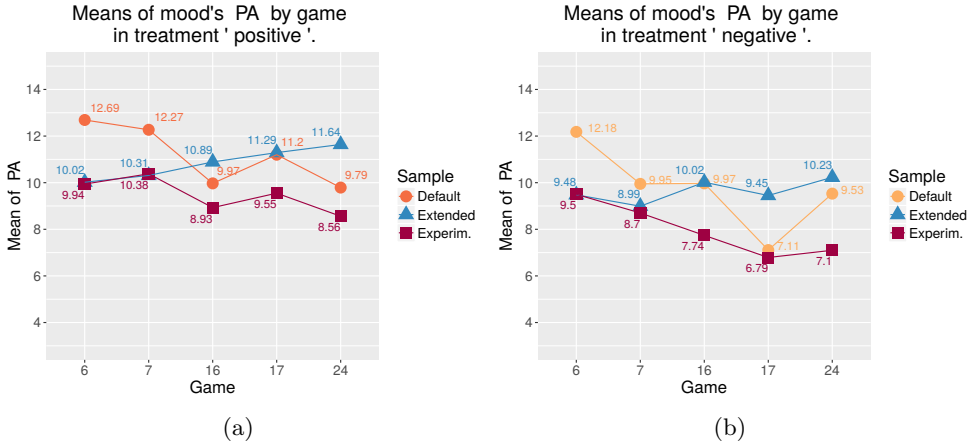


Figure 6.14: Means of  $PA$  for agents/participants in a positive (a), and negative treatment (b) in games 6, 7, 16, 17 and 24.

function updates the agent's mood following the idea that, a particular mood gets more intense the more experiences of this kind an individual has. This means that, if the agent has many negative experiences, it will tend to be in a negative mood more frequently than in a positive mood. This effect is mitigated in *GenIA*<sup>3</sup> by the process in charge of controlling the *affect temporal dynamics*, which needs to be adjusted in order to obtain the desired behavior.

In general, both the default and the extended simulations, represented the tendency to decrease of the number of agents with a negative mood in the positive treatment (Figure 6.12(b)), and the tendency to decrease of the number of agents with a positive mood in the negative treatment (Figure 6.13(a)). There are small variations from this behavior that can be explained because of the diffuse frontier between the positive and the neutral mood in the decision tree that has been used for mood classification.

Mood variations can be observed in more detail in Figures 6.14 and 6.15. They show the values of the means of the mood components  $PA$  (Figure 6.14) and  $NA$  (Figure 6.15) for games 6, 7, 16, 17, and 24 in the positive and negative treatments. Both the default and extended simulations properly represent the data of the experiment in that agents in the positive treatment are happier than

## 6. GENIA<sup>3</sup>'S AGENTS PLAYING BLACK JACK

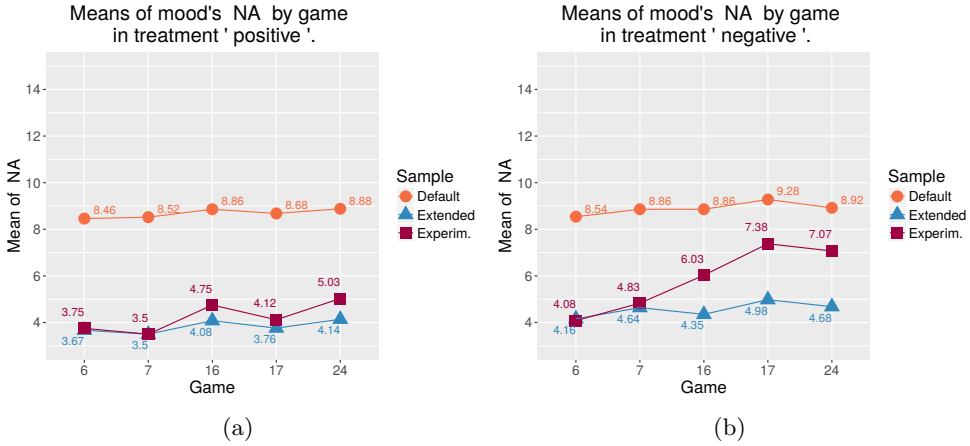


Figure 6.15: Means of  $NA$  for agents/participants in a positive (a), and negative treatment (b) in games 6, 7, 16, 17 and 24.

agents in the negative treatment, and that agents in the negative treatment are sadder than agents in the positive treatment after games 6 and 16. This means that the means of  $PA$  in the positive treatment are higher than the means of  $PA$  in the negative treatment, and that the means of  $NA$  in the negative treatment are higher than the means of  $NA$  in the positive treatment. However the main difference between the default and the extended simulation lies in the magnitude of the variations of  $PA$  and  $NA$ . In the default simulation the magnitude of the mood variations are higher than in the extended simulation. This can be explained because the default and the extended simulations use different methods for updating the mood.

### 6.3.2 Agent's Decisions

In order to compare the agent's plays with the participant's plays the percentage of the agents that performed the 'right' plays, according to the basic strategy was determined. Figure 6.16 shows these results for each of the three kinds of simulation (with rational agents, default and extended simulations), as well as the data of the experiment. Considering that the rational agents always play the basic strategy, their percentage of 'right' plays has always the maximum

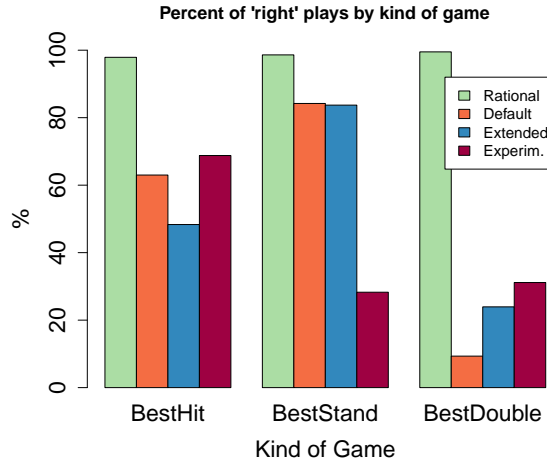


Figure 6.16: Percent of plays that match with the play suggested by the basic strategy in the non-training games.

value on each kind of game. The three possible kinds of game are those where the basic strategy suggests to hit ('best hit'), those where it suggests to double ('best double'), and those where it suggests to stand ('best stand'). This is determined according to the combinations of the player's and bank's cards (see Figure 6.3(b)). Both the default and the extended simulations share similar results, meaning that, the percentage of agents that hit in the 'best hit' games and the percentage of agents the double in the 'best double' games are more or less similar to these percentages in the data of the experiment. Also neither of both got a good approximation to to the percentage of participants that stand in the 'best stand' games. In the experiment, the participants tended to stand less and to take more risks (either doubling or hitting) than the agents on the simulation.

The percentage of kind of first plays performed on each simulation in game 7 is shown in Figure 6.17. The possible first plays are 'stand', 'hit', and 'double'. An additional category 'None' has also been included to denote those games where the player didn't need to perform any play (i.e., where either he or the bank got a Black Jack with the initial cards). Game 7 has been selected, because, at this point, the same amount of agents and participants have been

## 6. GENIA<sup>3</sup>'S AGENTS PLAYING BLACK JACK

---

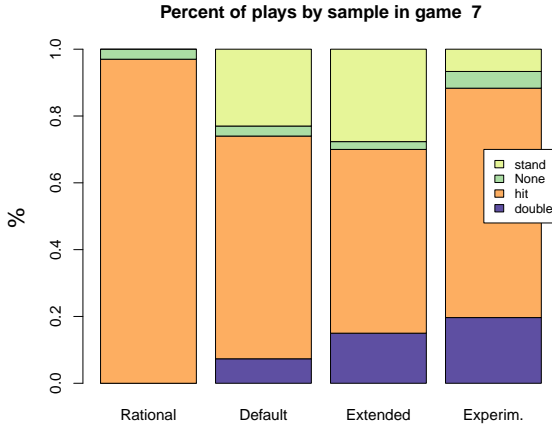


Figure 6.17: Percent of the first plays ‘double’, ‘hit’, and ‘stand’ for each kind of simulation in game 7 for the whole sample.

exposed to the same experiences, which is to win or to lose the previous game (game 6). As shown in Figure 6.17, the vast majority of rational agents play to hit, which is the ‘right’ play to do in game 7 according to what the basic strategy suggests. Also, the default and extended simulations share similar results regarding the proportion of plays performed in game 7 with small differences. Specifically, the percentage of agents that ‘hit’ a card in the first play of the default simulation (66.7% of ‘hit’ plays) is more close to the one of the experiment (68.7% of ‘hit’ plays) than to that of the extended simulation (55.0% of ‘hit’ plays). The same applies the percentage of ‘double’ plays of the extended simulation (15.0% of ‘double’ plays), which is more close to the one of the experiment (19.7% of ‘double’ plays) than the one of the default simulation (7.3% of ‘double’ plays). In general, the decision that has been better represented in the first play of game 7 is the decision of ‘hit’ in the default design.

When grouping agents (or participants) by their kind of mood in game 7 (see Figure 6.18), some other results can be observed. In Figures 6.18(a) and 6.18(b) it can be observed that the percent of participants with a positive mood that double the bet in game 7 (22.7% of ‘double’ plays) is higher than

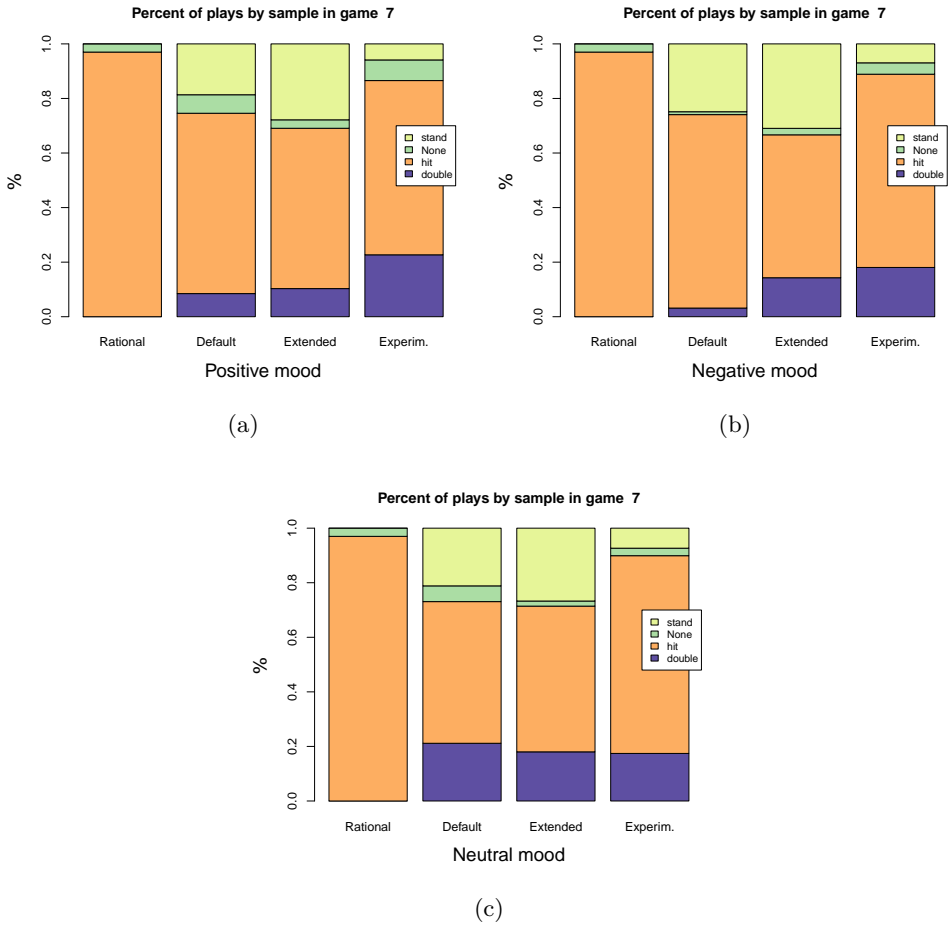


Figure 6.18: Percent of the first plays ‘double’, ‘hit’, and ‘stand’ for each kind of simulation in game 7 by kind of mood.

## EVALUATION

---

the percent of participants with a negative mood that double the bet in game 7 (18.1% of ‘double’ plays). This tendency has been represented in the default simulation (with values of 8.5% and 3.2% respectively) but not in the extended simulation. There is another related result that can be observed in the data of the experiment, which is that participants in a negative mood tend take more risks, by hitting a new card, than the participants in the positive mood. Specifically 70% of participants hit a new card in the first play when they were in a negative mood, and 63.9% of participants hit a new card in the first play when they were in a positive mood. In the default simulation these percentages are 70.9% and 66.1% respectively. Nevertheless, regarding to the agents and participants in a neutral mood, better approximations were obtained by the extended design, whose percentages of agents that play ‘hit’ and ‘double’ (53.4% and 18.0% respectively) are closer to the percentages of the experiment (72.5% and 17.4% respective) than the percentages of the default simulation (51.9% and 21.2% respectively). In general, the behavior better represented between both the default and the extended simulations is that of the agents with the positive mood in the default simulation. Also, according to the results shown in Figure 6.17, the participant’s behavior was better represented by the default simulation.

## Conclusions

In Chapter 5 two classical games of experimental economics were used, in order to address social preferences, by endowing players with a ‘human’ component. During each game, one of the participant players experienced changes on its affective state according to what happened in the game, and in turn, the player’s decisions were influenced by its affective state. Besides, Chapter 6 describes the design and execution of an experiment with 300 real participants, where participant’s mood was induced, and where participants’ decisions were recorded with the aim of determining the participants’ mood influence on their decisions. With the execution of this experiment, the Goal G4 of this thesis has been achieved. These examples and case study helped to verify and validate some aspects of the *GenIA*<sup>3</sup> architecture and platform stated at the beginning



of the ‘Evaluation’ Part of this thesis.

The classical games ‘prisoner’s dilemma’ and ‘trust game’ were used in order to validate whether *GenIA*<sup>3</sup> was able to solve typical issues of social preferences that are traditionally solved by classical games. In both examples, one of the participant players was endowed with affective characteristics and with a personality. In both cases a rational solution and also the solution with affective agents was proposed. In the prisoner’s dilemma it has been illustrated the way in which good strategies such as the Tit-For-Tat strategy can be represented by an affective agent, and how the behavior of this agent can be customized by changing its rationality level. This facility is also illustrated in the trust game approach, which used results with real participants that included the participants’ personality. By changing the values of the agent’s personality traits, it experienced different levels of anger, and thus different decisions (amount of money sent) were made. Besides, by modifying the agent’s rationality level, the influence of its affective state on its decisions changed accordingly. The way that these approaches address social preferences is more in line with human behavior than traditional methods. Firstly, because they were able to get results similar to those observed in human participants; and secondly because they consider affective characteristics and personality, which have been extensively proven to be determinant when modeling human behavior. With these examples we partially fulfill the Goal G5.3 of this thesis.

The case study that has been used in Chapter 6 has also helped to reach other relevant conclusions. As part of this case study some issues of *GenIA*<sup>3</sup> were validated. They include the extent to which agents experience the participants’ mood in similar circumstances, the extent to which their decisions were similar to the participant’s decisions, whether the default design of *GenIA*<sup>3</sup> was able to simulate the results of the experiment, and whether an extended design was able to improve these results. By carrying out these validations, the Goals G5.1 and G5.2 of this thesis have been achieved. Three different simulations were performed in order to perform the proposed validations. A simulation with fully rational agents (rational simulation), a simulation of agents following the default design of *GenIA*<sup>3</sup> (default simulation), and a simulation of agents following an extended design of *GenIA*<sup>3</sup> (extended simulation). Agents

## EVALUATION

---

that were completely rational didn't have affective characteristics, and they played always according to what the Peter Griffin's basic strategy suggests [Griffin, 1996]. Regarding the agents' decisions, in both the default and the extended simulations, agent's behavior was more similar to the participant's behavior than the behavior of rational agents, what shows that *GenIA*<sup>3</sup>'s affective agents better represent human like behaviors than merely rational agents. Nevertheless, the percentage of agents by kind of decision in the default and extended simulations is similar, having the default simulation slightly better results than the extended simulation. This result demonstrates that the *GenIA*<sup>3</sup>'s default design is generic enough to properly represent agent's affective characteristics and behavior without the need for additional extensions. Nevertheless, regarding the agent's mood, although both the default and extended simulations properly simulate the mood tendencies of participants when they won or lost a game (i.e., after the game they were happier than before the game when they won, and sadder when they lost), some differences were found. In most cases (mainly in the positive treatment), the variation of the mood components (*PA* and *NA*) wasn't as close to the values of the experiment as expected. This is understandable considering that, mood update specific parameters may vary to a greater or lesser extent depending on the application domain. In order to address this issue, regression models were fitted so that better predictions of mood variations were made. These models were included in the appraisal and mood update processes of the extended simulation. This has demonstrated that *GenIA*<sup>3</sup> is versatile enough to be extended and adapted when required, in order to better represent human behavior.

Finally, in Chapters 5 and 6, it has been demonstrated how *GenIA*<sup>3</sup> offers proper mechanisms to implement the influence of the affective state in decisions. One of those mechanisms is through the categorization of affective states, which allows to implement different behaviors depending on the agent's affective state (that evolves and changes over time) at each time point. The other mechanism is by recovering the agent's affective state exact value, and by using this value as a parameter of subsequent decisions or processing. Besides, the examples and case study used in Chapters 5 and 6, have shown the flexibility of *GenIA*<sup>3</sup> in all

cases, since it has been easily extended and adapted to fulfill the requirements of each case. The examples and case study have also shown that agent's rational strategies can be improved with *GenIA*<sup>3</sup> in order to get more human like behaviors, thus fulfilling the Goal G5.3 of this thesis.



Part V

**Conclusions**



---

## Conclusions and Future Work

This chapter summarizes the main conclusions of this thesis work, and it presents a global vision of the thesis' main contributions. The future lines of research that are derived from this thesis are also discussed, and finally related publications are presented as well.

### 7.1 Conclusions

As stated in Chapter 1, the main goal of this work has been to offer an approach where basic theoretical and practical elements are integrated in a toolbox for conceiving affective and intelligent agents. This goal has been fulfilled in that useful related tools have resulted from the current research. These tools include: (1) a theoretical framework that includes conceptual systems of the main concepts and processes that a generic computational approach should consider in order to model affective intelligent agents; (2) the general purpose affective agent architecture *GenIA*<sup>3</sup>, on the basis of the proposed theoretical framework; (3) a default design for *GenIA*<sup>3</sup>; (4) an extension of the Jason main reasoning cycle, and of the Jason language, in order to build agents compliant with *GenIA*<sup>3</sup>; (5) an extension of the Jason platform, in such a way, that the agents can be executed with *GenIA*<sup>3</sup>'s default design as well as with extended

## 7. CONCLUSIONS AND FUTURE WORK

---

designs; (6) and data of an experiment with real participants were obtained, which have helped to compare *GenIA*<sup>3</sup>'s agents behaviors with human beings' behaviors.

The scenario of the above mentioned experiment was successfully simulated in the *GenIA*<sup>3</sup>'s platform, and the participants' behavior was properly reproduced. The same amount of agents (under the same experimental conditions of the participants), experienced similar affective states, and had similar tendencies on the variation of the affective state. This confirms part of the hypothesis stated in Chapter 1, because the proposed architecture was able to simulate the human behavior in a real situation. The improvement of human-like simulations with respect to simulations of completely rational agents, was another part of the hypothesis, which has been confirmed in three different case studies. For each case study, situations where agents offered rational solutions were exposed, and more plausible behaviors were obtained with *GenIA*<sup>3</sup>'s affective agents. For example, when simulating the classical game 'prisoner's dilemma', a similar strategy to Tit-For-Tat (which is the strategy that won twice Axelrod's tournament [Axelrod & Hamilton, 1981]) was developed in an affective agent, which was able to affectively react to the opponent's moves, and to behave under the influence of its affective state.

In the hypothesis of this thesis, it was also stated that the proposed architecture would allow the incremental development of its individual components. This has been confirmed for several components. For example, the agent's affective state has been extended for it to be compliant with the PANAS model of affect used in the experiment simulation. The agent's affective state has also been extended with the level of 'anger' used in the Trust Game, in substitution of the PAD model of affect that is part of the default design. Same way, the agent's personality and appraisal variables could be customized, in order to be adapted to the specific requirements of each case study. Besides, the process of initialization of the affective state could successfully be customized for initializing the affective state (and, hence, for setting the affective state's 'equilibrium' value), by either using personality traits in the trust game, or by using specific values previously saved on a data base in the experiment simulation. The processes of appraisal and updating of the affective state could



also be successfully extended in order to fulfilled the requirements of each case study. Specifically, in the experiment simulation, although the results obtained with agents that had the *GenIA*<sup>3</sup>'s default design (default simulation) were satisfactory, the simulation with agents that had an extended design (extended simulation) improved the results of the default simulation. By extending the processes of appraisal and updating of the affective state, in accordance with the models fitted for the experiment's data, the proportion of agents that experienced affective states similar to those of the participants of the experiment was higher than the proportion of the default simulation.

### 7.1.1 Contributions

Considering the great difficulty of integrating knowledge of multiple disciplines in such a way that they can reuse one another's contributions, the main global contribution of this work was to favor the interdisciplinary exchange, the intradisciplinary coordination, and the incremental research. This has been possible thanks to the support of specialists on psychology, economy, affective computing, machine learning, and computer science in general. On the other hand, several contributions are derived from this research at various levels, in accordance to the objectives proposed in Chapter 1. These contributions include:

On the *State of the Art Level*, it has been offered a horizontal review of the main concepts and processes that are traditionally addressed, mainly in psychology, but also in neurology. These concepts and processes are related to the way the affective state changes on individuals, and how it influences cognition and practical reasoning. This review has been performed with a computational perspective, with the aim of gathering, in a formal and computable representation, the main concepts and processes that a rational agent with an affective component should possess.

A **generic theoretical framework** was proposed along with *two conceptual systems* of the main agent attributes and processes for BDI affective agents, which are grounded in the main psychological theories

## 7. CONCLUSIONS AND FUTURE WORK

---

on emotions, mood, and personality. This framework and conceptual systems allow to conceive affective agents with practical reasoning that live in an agent community. They also allow to compare affective agents approaches regarding two aspects: (a) the extent to which these approaches cover the core concepts and processes that the theoretical framework and conceptual systems propose, and (b) the extent to which the practical reasoning and affect-related issues are integrated into affective agents approaches.

On the *Agent Level*, a general purpose architecture for rational and affective agents (*GenIA*<sup>3</sup>) has been proposed. This architecture is in line with the “theoretical toolbox of basic theory-elements, formulated in a common language, from which theories of emotional agents (or of emotion modules for agents) can be constructed” suggested by Reisenzein *et al.* [2013], unlike approaches such as FAtiMA Modular [Dias *et al.*, 2014], which is more in line with a “modularized architecture”. One of the advantages of an architecture like *GenIA*<sup>3</sup> is that, it has a completely functional **default design** (which is another contribution of this thesis), that is ready to use, that can be extended only if the results with the default design are not satisfactory enough, and where structures for the main affect related processes are already part of the basic structure (such a coping structure and a beliefs modulator structure). In FAtiMA Modular, for example, an agent that has only the core structure can’t do anything by default, and extra components must be included in order to adapt the architecture to the requirements. Besides, *GenIA*<sup>3</sup> allows to set different relations between affectivity and rationality, by either prioritizing ‘affective plans’ or ‘rational plans’, or by setting different synchronization relations between affective and rational cycles. Another advantage of *GenIA*<sup>3</sup>’s agents is that they can be implemented in an agent language adapted to the *GenIA*<sup>3</sup>’s architecture.

The **extension of the AgentSpeak’s operational semantics in a semi formal way and of the Jason language** has been another contribution of this thesis. Such agent language allows to create affective

BDI agents by using the main attributes proposed on the conceptual system of Chapter 2, in a compact and elegant notation and avoiding ambiguities. These attributes include the agent’s coping strategies, personality traits, rationality level, and concerns, among others. The agent language also includes other useful mechanisms such as a direct access to the values of the agent’s affective state and personality, through specific structures called *internal actions* (see [Bordini *et al.*, 2007] for a reference), and plans categorization according to one or more affective categories for them to be executed for a given affective state.

On the *System Level*, a **software a platform** for executing *GenIA*<sup>3</sup>’s agents has also resulted of this thesis work. This platform is an extension of the Jason platform, which specifically extends the ‘Centralized’ infrastructure of Jason. This extension offers structures that allow to execute *GenIA*<sup>3</sup> based agent societies, with different levels of customization. At the agent level, it offers mechanisms for customizing the agents’ definition through proper interfaces, as well as the synchronization of the agent’s affective and reasoning cycles. At the system level, the affective categories that are used for the affective selection of plans can also be defined. Such generic structures allow to define MAS of diverse nature and agents’ behaviors.

On the *Applications Level*, the generic nature of this approach makes possible to create agents in **different application domains**. For example, for computational narratology<sup>1</sup>, affective agents could be created that represent the way humans affectively react to environmental circumstances, and the way this affective state influences subsequent decisions. Moreover, an embodied virtual character, which expresses emotions through facial expressions, can use an adapted model of coping strategies, in order to properly orchestrate facial expressions for creating believable behaviors. On the other hand, mixed application domains could be used

---

<sup>1</sup>According to Mani [2014], it can be defined as “The study of narrative from the point of view of computation and information processing (...). Its scope includes the approaches to storytelling in artificial intelligence systems and computer (and video) games, the automatic interpretation and generation of stories, and the exploration and testing of literary hypotheses through mining of narrative structure from corpora”.

## 7. CONCLUSIONS AND FUTURE WORK

---

as well. For example, a virtual character able to play strategic games, may need expressive behaviors, mechanisms for affectively appraising situations, and mechanisms for making rational decisions, that are possibly influenced by its affective state. Additionally, by being part of a MAS platform, *GenIA*<sup>3</sup>'s agents are able to simulate human societies where agents can negotiate, argue, work in normative environments, be part of organizations, and, in general, be able to use all the tools that agreement technologies provide<sup>1</sup>. Finally, other possible applications, are derived from the multidisciplinary nature of this thesis work as well. For example, for social psychology or experimental economics, groups of artificial human-like entities could be used for experimentation in the safety of a virtual world, without the need of physical infrastructures or related logistics, and without the restrictions that, for example ethical issues may impose.

### 7.2 Future Work

This PhD thesis has a multidisciplinary and hybrid nature, thus it opens a wide range of research lines that can advance the current research. These research lines include: the study of alternative default mechanisms for *GenIA*<sup>3</sup>, the extension of Jason's capabilities for working with prospective emotions, the improvement of *GenIA*<sup>3</sup>'s mechanisms to prioritize plans, the creation of structures for reusing the affective reasoning "machinery", the automatic evaluation of a situation's "complexity" for adjusting rationality levels, and the implementation of new application domains in *GenIA*<sup>3</sup>. Next, these possible research lines are described in more detail.

**Study of alternative default mechanisms for *GenIA*<sup>3</sup>.** The default mechanism that has been used in this work as the *GenIA*<sup>3</sup>'s default design, has been selected on the basis of widely accepted psychological theories. Nevertheless, other theories could be used as well. It would be interesting

---

<sup>1</sup>See, for example, <http://jason.sourceforge.net/wp/projects/> for a reference of Jason's related projects that include agreement technologies.

to investigate how these theories could be used as alternatives to the *GenIA*<sup>3</sup>'s default design. These alternatives could be integrated in such a way that, programmers could “switch” between different designs, which could be part of the affective platform. This is an approach that can be carried out in a short-term, since *GenIA*<sup>3</sup>'s platform inner structure already supports different kinds of extensions for customizing the agent behavior.

**Extension of Jason’s capabilities for working with prospective emotions.** The possibility of experiencing prospective emotions (those that result from the appraisal of future events) is an important human characteristic. Due to the generic nature of *GenIA*<sup>3</sup>, it would be interesting to find out what default mechanisms can be included in *GenIA*<sup>3</sup> in order to better experience prospective emotions. Some of the main ingredients of prospective emotions are expectations and beliefs probabilities. Currently *GenIA*<sup>3</sup> offers a default mechanism for reasoning about what we called Jason beliefs’ probabilities (that we interpret as expectations). Nevertheless, mechanisms for discriminating between the probability of propositions’ occurrence and expectations are required. By considering the probability of propositions’ occurrence it would be also necessary to either create new (or integrate existing) mechanisms for probabilistic reasoning. Related to this, we are currently involved in the refinement of *GenIA*<sup>3</sup>'s mechanisms for reasoning about expectations [Taverner *et al.*, 2016].

**Improvement of *GenIA*<sup>3</sup>'s mechanisms to prioritize plans.** The default mechanism of *GenIA*<sup>3</sup> to prioritize plans works with the agent’s rationality level, giving a higher priority to the affective plans when the rationality level is low, and giving a higher priority to the non-affective plans when the rationality level is high. An interesting contribution would be to offer the possibility of setting priorities to affective plans over other affective plans and to non-affective plans over other non-affective plans. This can be done by using either Jason’s own structures or by creating new ones. Moreover, considering that the agent’s personality

## 7. CONCLUSIONS AND FUTURE WORK

---

(as well as the affective state) has an important influence on behavior, it would also be of interest to include personality profiles, in such a way that these profiles also determine what plans may be selected.

**Automatic evaluation of a situation “complexity”.** The agent attribute ‘rationality level’ proposed in this work is used to select, to a greater or lesser extent, affective plans (i.e., to carry out, to a greater or lesser extent, affective decisions). On the other hand the “complexity” of a situation determines the extend to which the affective state may influence decisions [Forgas, 1995]. By including a default mechanism that determines how “complex” a situation is on *GenIA*<sup>3</sup>, it wouldn’t be necessary to fix a numerical value for the agent’s rationality level, since it could be automatically determined. This could be done by establishing an inverse relation between the rationality level and the situation “complexity”.

**Creation of structures for reusing the *GenIA*<sup>3</sup>’s reasoning “machinery”.** One of the humans’ capabilities consists of reasoning about other’s mental states and emotions. An interesting line of research would be to study how to use *GenIA*<sup>3</sup>’s affective (and non-affective) reasoning “machinery” as a service. This will allow to anticipate other agents’ affective state [Alfonso *et al.*, 2015a].

**Implementation of new application domains in *GenIA*<sup>3</sup>.** Finally, in order to explore the limits of *GenIA*<sup>3</sup>, it would be interesting to use case studies of new application domains that allow to find out whether new default mechanisms could be assigned to other components of *GenIA*<sup>3</sup>.

### 7.3 Related Publications

#### Journals

Alfonso, B., Vivancos, E. & Botti, V. (2017). Toward Formal Modeling of Affective Agents in a BDI Architecture. *ACM Transactions on Internet Technology (TOIT)*, 17, Article 5. **JCR 0.705 - 71/106 Q3 T2**

Alfonso, B., Vivancos, E. & Botti, V. (2017). From Affect Theoretical Foundations to Computational Models of Intelligent Affective Agents. *Manuscript submitted for publication.*

Alfonso, B., Botti, V., Garrido, A. & Giret, A. (2014). A MAS-Based Infrastructure for Negotiation and its Application to a Water-Right Market. *Information Systems Frontiers*, 16, 183–199. **JCR 1.077 - 63/139 Q2 T2**

#### Conferences and Workshops

Alfonso, B., del Val, E. & Alberola, J.M. (2016). An Agent Model of Student’s Affect for Adapting Schooling Strategies. In *Methodologies and Intelligent Systems for Technology Enhanced Learning, 6th International Conference*, 15–22, Springer.

Taverner, J., Alfonso, B., Vivancos, E. & Botti, V. (2016). Integrating Expectations into Jason for Appraisal in Emotion Modeling. In *Proceedings of the 8th International Conference on Evolutionary Computation Theory and Applications (ECTA). Part of the 8th International Joint Conference on Computational Intelligence (IJCCI)*, vol. 1, 231–238.

Alfonso, B., Pynadath, D.V., Lhommet, M. & Marsella, S. (2015). Emotional perception for updating agents’ beliefs. In *International Conference on Affective Computing and Intelligent Interaction (ACII)*, 201–207, IEEE. **Core C**

Alfonso, B., Vivancos, E., Botti, V. & Hernández, P. (2015). Building Emotional Agents for Strategic Decision-Making. In *Proceedings of the 7th International Conference on Agents and Artificial Intelligence (ICAART)*, 390–397.

Alfonso, B., Vivancos, E. & Botti, V.J. (2014). An Open Architecture for Affective Traits in a BDI Agent. In *Proceedings of the 6th International Conference on Evolutionary Computation Theory and Applications*

## 7. CONCLUSIONS AND FUTURE WORK

---

(ECTA). *Part of the 6th International Joint Conference on Computational Intelligence (IJCCI)*, 320–325.

Alfonso, B., Vivancos, E., Botti, V. & García-Fornes, A. (2011). Integrating Jason in a Multi-Agent Platform with Support for Interaction Protocols. In *Proceedings of the compilation of the co-located workshops on AGERE!'11 , Conference on Systems, Programming, Languages and Applications: Software for Humanity (SPLASH)*, 221–226, ACM, NY, USA.



---

---

## Appendix A: Ethical discussion

### A.1 Ethical discussion

When reflecting on the proposal performed on this thesis work, several ethical issues may come to mind. In [Scheutz, 2012], Matthias Scheutz discusses the viability, from an ethical perspective, of building affective artificial agents, by highlighting the “affect dilemma”. The affect dilemma focuses on the idea that the very fact of creating artificial agents that can potentially suffer can make us to question ourselves whether creating such agents is morally acceptable. But, on the other hand, considering that human-computer interaction is nowadays part of human life, and that the human being has a fully affective nature, creating non-affective artificial agents could cause harm and not effective human-computer interactions. On the other hand, public opinion often has concerns that involve moral principles such as the unnaturalness of artificial agents that seem to have emotions, or the morally unacceptable fact of venturing in part of existence [Cowie, 2015]. Nevertheless, there are reasons for investigating on affective processes in artificial systems that make the utility of artificial agents overcome possible dangers resulting from them. According to [Scheutz, 2012] these reasons include, among others, the possibility of verifying grounding theories that computational models of affect offer; artificial agents can help to make ethical decisions in a way that, for example, consequences that produce negative affects in human beings could be avoided; or, in general, for preventing human suffering as a consequence of unidirectional emotional bonds with social artificial agents or robots, which can be avoided with a bidirectional emotional exchange.

## APPX A. ETHICAL DISCUSSION

---

Our approach has been conceived to foster the above mentioned reasons for which artificial affective agents should be designed, and every approach based in ours should be designed with the same purposes. In order to avoid any misuse of the research performed on this thesis work, every approach based in ours should, at least, fulfill the main ethical principles proposed by the UK Research Council for Engineering and Physical Science (EPSRC) [Boden *et al.*, 2017]. These principles refer to research on robotics but they can be applied to artificial agents too [Cowie, 2015]. These principles state that, for example, in this kind of affective artificial systems, “the illusion of emotions and intent should not be used to exploit vulnerable users”.

On the other hand, considering the general nature of our approach, several additional ethical issues may arise by using our platform in specific application domains. For example, in the case of affective systems that are designed to be companions it is important to take into account risks such as those exposed in [Turkle, 2010] (e.g., that of sidestepping “encounters with friends and family” as a consequence of the interaction with artificial systems). Or, in the case of affective systems in medicine, the potential extreme consequences if therapies with artificial affective systems offer results that are contrary to those expected. It is a responsibility of researchers to assess what ethical principles can be violated with each particular approach derived from ours.

The human participation in our research also required several ethical issues to be considered. These considerations have been done according to the American Psychological Association (APA) code of ethics<sup>1</sup>. Specifically, the experiment was conducted by the LINEEX<sup>2</sup>, which is an specialized experimentation center that guarantees the fulfillment of these ethical principles and of other technological codes such as the maintenance of the privacy and integrity of the participants’ personal data. Besides, participants were offered inducements for their participation, truthful information was offered to the participants regarding the experiment, and each session of the experiment was ended with a debriefing stage.

---

<sup>1</sup><http://www.apa.org/ethics/code/index.aspx>

<sup>2</sup><https://lineex.es>

---

---

## Appendix B: Affective Agents' Grammar and Examples

This appendix contains the grammars of the agent language and Multi-Agent Systems language that were used to generate the parsers used in our Jason extended version. This appendix also contains, in two different sections, the main code for the 'Prisoner's dilemma' (PD) and the 'Trust game' (TG) simulations respectively. The code was written in our Jason extended version and both examples show the main code of the participant agents (see Figures B.4.2, B.4.3, B.4.4, B.4.5, B.5.6, and B.5.7). Additionally, in the PD example, the code of the MAS configuration is also shown (see Figure B.4.1).

### B.2 Syntactic Specification of the Agent Language

```
agent ::= ( directive )* ( belief ( directive )* )* ( concerns
( directive )* )? ( personality ( directive )* )? (
initial_goal ( directive )* )*( plan ( belief )* ( directive
)* )* <EOF>

directive ::= "{" ( <TK_BEGIN> pred "}" agent | pred "}" )

belief ::= literal_prob ( ":-" log_expr )? "."

initial_goal ::= "!" literal "."

plan ::= ( <TK_LABEL_AT> literal_prob )? trigger ( ":" log_expr
) ? ( "<-" plan_body )? "."

trigger ::= ( "+" | "-" | "^" ) ( ( "!" | "?" ) )? literal_prob

plan_body ::= plan_body_term ( ";" )? ( plan_body )?

plan_body_term ::= plan_body_factor ( <TK_POR> plan_body_term )?
```

## APPX B. AGENTS' GRAMMAR AND EXAMPLES

---

```

plan_body_factor ::= ( stmtIF | stmtFOR | stmtWHILE | body_formula )
                    ( <TK_PAND> plan_body_factor )?
stmtIF           ::= <TK_IF> “ ( ” log_expr “ ) ” rule_plan_term
                    ( ( <TK_ELSE> )? rule_plan_term )?
stmtFOR        ::= <TK_FOR> “ ( ” log_expr “ ) ” rule_plan_term
stmtWHILE     ::= <TK_WHILE> “ ( ” log_expr “ ) ” rule_plan_term
body_formula  ::= ( “!” | “!!” | “?” | ( “+” ( ( “+” | “<” | “>” ) )? ) | ( “-” ( “+”
                    )? ) )? ( rule_plan_term | log_expr )
rule_plan_term ::= “{” ( ( <TK_LABEL_AT> pred )? trigger ( “:” log_expr
                    )? ( ( “<-” | “;” ) )? ( literal_prob “:-” log_expr )? (
                    plan_body )? “}”
literal       ::= ( ( ( ( <ATOM> | var ) )? “:” )? ( <TK_NEG> )? ( pred |
                    var ) ) | <TK_TRUE> | <TK_FALSE> )
literal_prob  ::= ( ( ( ( <ATOM> | var ) )? “:” )? ( <TK_NEG> )? ( ( pred
                    ( range )? ) | var ) ) | <TK_TRUE> | <TK_FALSE> )
pred         ::= ( <ATOM> | <TK_BEGIN> | <TK_END> ) ( “ ( ” terms “ ) ” )?
                    ( list )?
terms        ::= term ( “ , ” term )*
term         ::= ( list | rule_plan_term | log_expr )
list        ::= “[” ( term_in_list ( “ , ” term_in_list )* ( “ | ” ( <VAR>
                    | <UNNAMEDVAR> | list ) )? )? “ ] ”
term_in_list ::= ( list | arithm_expr | string | rule_plan_term | prob )
log_expr     ::= log_expr_trm ( “ | ” log_expr )?
log_expr_trm ::= log_expr_factor ( “ & ” log_expr_trm )?
log_expr_factor ::= ( <TK_NOT> log_expr_factor | rel_expr )
rel_expr     ::= ( arithm_expr | string ) ( ( “<” | “<=” | “>” | “>=” | “==”
                    | “\==” | “=” | “=. .” ) ( arithm_expr | string | list |
                    rule_plan_term ) )?
arithm_expr  ::= arithm_expr_trm ( ( “+” | “-” ) arithm_expr_trm )*
arithm_expr_trm ::= arithm_expr_factor ( ( “*” | “/” | <TK_INTDIV> |
                    <TK_INTMOD> ) arithm_expr_factor )*
arithm_expr_factor ::= arithm_expr_simple ( ( “**” ) arithm_expr_factor )?
arithm_expr_simple ::= ( <NUMBER> | “-” arithm_expr_simple |
                    “+” arithm_expr_simple | “ ( ” log_expr “ ) ” | function )
function     ::= literal_prob
var         ::= ( <VAR> | <UNNAMEDVAR> ) ( list )?
prob        ::= <TK_PROB> “ ( ” <NUMBER> “ ( “ , ” ( <TK_POSITIVE> |
                    <TK_NEGATIVE> ) )? “ ) ”
string      ::= <STRING>
personality ::= <TK_PERSONALITY> “:” “{” listNumbers
                    ( “ , ” <NUMBER> )? ( “ , ” copingst )? “}” “.”
listNumbers ::= “[” ( <NUMBER> ) ( “ , ” <NUMBER> )* “ ] ”
concerns   ::= <TK_CONCERNS> “ ( ” <VAR> “ ) ” “:-” log_expr “.”
copingst   ::= <TK_COPINGST> “:” ( cs )+
cs         ::= <TK_CS> “ ( ” log_expr “ , ” <ATOM> “ ) ”
                    “->” plan_body “.”
range      ::= “<” arithm_expr “:” arithm_expr “>”

```

### Tokens

```
SKIP : {
  " "
|  "\t"
|  "\n"
|  "\r"
|  <"/" (~["\n","\r"])* ("\n" | "\r" | "\r\n")>
|  <"/*" (~["*"])* "*" ("*" | ~["*","/"] (~["*"])* "*" )* "/">
}

TOKEN : {
  <VAR : (<UP_LETTER> (<CHAR>)* ) >
}

TOKEN : {
// Predefined
  <TK_TRUE: "true">
| <TK_FALSE: "false">
| <TK_NOT: "not">
| <TK_NEG: "~">
| <TK_INTDIV: "div">
| <TK_INTMOD: "mod">
| <TK_BEGIN: "begin" >
| <TK_END: "end" >
| <TK_LABEL_AT: "@"> // special chars

| <TK_IF: "if" >
| <TK_ELSE: "else" >
| <TK_FOR: "for" >
| <TK_WHILE: "while" >

| <TK_PAND: "|&|" >
| <TK_POR: "|||" >

| <TK_PERSONALITY: "personality__" >
| <TK_CONCERNS: "concerns__" >
| <TK_COPINGST: "copingst__" >
| <TK_CS: "cs__" >
| <TK_PROB: "prob__" >

// Numbers
| <NUMBER: ["0"- "9"] (["0"- "9"])*
  | (["0"- "9"])* "." (["0"- "9"])+ (["e", "E"] (["+","-"])?
  (["0"- "9"])+)?
  | (["0"- "9"])+ (["e", "E"] (["+","-"])? (["0"- "9"])+) >
}
```

## APPX B. AGENTS' GRAMMAR AND EXAMPLES

---

```
// Strings
| <STRING: "\" ( ~["\\", "\\", "\n", "\r"]
  | "\\\" ( ["n", "t", "b", "r", "f", "\\", "\', "\""]
  | ["0"- "7"] (["0"- "7"])?
  | ["0"- "3"] ["0"- "7"] ["0"- "7"]))) * "\">

// Identifiers
| <ATOM      : (<LC_LETTER> | "." <CHAR>) (<CHAR> | "." <CHAR>)*>
| <UNNAMEDVAR: ("_" (<CHAR>)* ) >
| <CHAR      : (<LETTER> | <DIGIT> | "_")>
| <LETTER    : ( <LC_LETTER> | <UP_LETTER> )>
| <LC_LETTER : ["a"- "z"]>
| <UP_LETTER : ["A"- "Z"]>
| <DIGIT    : ["0"- "9"]>
}
```

### B.3 Syntactic Specification of the Multi-Agent System Language

```
mas ::= <MAS> <ASID> "{ infra environment control agents directives classpath sourcepath affectCategories }"

infra ::= ( <INFRA> ":" classDef )?
agents ::= ( <AGS> ":" ( agent )+ )?
agent ::= <ASID> ( fileName )? ASoptions ( <ASAGARCHCLASS> classDef | <ASAGCLASS> classDef | <BBCCLASS> classDef | <PCLASS> classDef | <ASCLASS> classDef | <EMENGINECLASS> classDef | "#" <NUMBER> | <AT> <STRING> ) * ";"

fileName ::= ( <PATH> )? <ASID> ( "." <ASID> )?
classDef ::= ( <ID> | <ASID> ) ( "." ( <ID> | <ASID> ) ) * ( "(" ( ( parameter ) ( "," ( parameter ) ) * )? ")" )?

parameter ::= ( classDef | <NUMBER> | <STRING> | listParameters )
listParameters ::= "[" ( parameter ) ( "," parameter ) * "]"

ASoptions ::= ( "[" procOption ( "," procOption ) * "]" )?
procOption ::= ( <ASOEE> "=" <ASOEEV> | <ASOIB> "=" <ASOIBV> | <ASOSYNC> "=" <ASOBOOL> | <ASONRC> "=" <NUMBER> | <ASOV> "=" <NUMBER> | <NACPRC> "=" <NUMBER> ":" <NUMBER> | <ASID> "=" ( <STRING> | <ASID> | <NUMBER> | <ID> | <ASOBOOL> ) )

environment ::= ( <ENV> ":" classDef ( <AT> <STRING> )? )?
control ::= ( <CONTROL> ":" classDef ( <AT> <STRING> )? )?
```

### B.3. Syntactic Specification of the Multi-Agent System Language

```
classpath ::= ( <CLASSPATH> ":" ( <STRING> ";" )+ )?
sourcepath ::= ( <SOURCEPATH> ":" ( <STRING> ";" )+ )?
directives ::= ( <DIRECTIVE> ":" ( <ASID> "=" classDef ";" )+ )?
affectCategories ::= ( <AFFCAT> ":" ( affectCategory ) ( "," affectCategory )* ";" )?
affectCategory ::= <ASID> " ( ( ( <NUMBER> ) | ( "-" <NUMBER> ) | range ) ( "," ( ( <NUMBER> ) | ( "-" <NUMBER> ) | range ) )* " )"
range ::= "<" ( ( <NUMBER> ) | ( "-" <NUMBER> ) ) ":" ( ( <NUMBER> ) | ( "-" <NUMBER> ) ) ">"
```

### Tokens

```
SKIP : {
  " "
|  "\t"
|  "\n"
|  "\r"
|  <"//"> (~["\n", "\r"])* (" \n" | " \r" | " \r\n" )>
|  <"/ *> (~["*"])* "*" ("*" | ~["*", "/" ] (~["*"])* "*" )* "/">
}
```

```
TOKEN : {
```

```
// Predefined
<MAS:      "MAS">
| <AGS:     "agents">
| <ENV:     "environment">
| <CONTROL: "executionControl">
| <AT:      "at">
| <INFRA:   "infrastructure" >
| <CLASSPATH: "classpath" >
| <SOURCEPATH: "aslSourcePath" >
| <AFFCAT:  "affcategories" >

| <ASOEE:   "events" >
| <ASOEEV:  ("discard"|"requeue"|"retrieve") >

| <ASOIB:   "intBels">
| <ASOIBV:  ("sameFocus"|"newFocus") >

| <ASONRC:  "nrcbp" >

| <ASOV:    "verbose" >

| <DIRECTIVE: "directives" >
```

## APPX B. AGENTS' GRAMMAR AND EXAMPLES

---

```
| <ASOSYNC:      "synchronised" >
| <ASOBOOL:      ("true" | "false" ) >

| <ASAGCLASS:    "agentClass" >
| <ASAGARCHCLASS: "agentArchClass" >
| <BBCLASS:      "beliefBaseClass" >
| <PCLASS:       "personalityClass" >
| <ASCLASS:      "affStateClass" >
| <EMENGINECLASS: "emEngineClass" >
| <NACPRC:       "nacprc" >

// Numbers
| <NUMBER: ["0"-"9"] (["0"-"9"])*
| ( ["0"-"9"])* "." ( ["0"-"9"])+ ( ["e","E"] ( ["+", "-"])?
| ( ["0"-"9"])+ )?
| ( ["0"-"9"])+ ( ["e","E"] ( ["+", "-"])? ( ["0"-"9"])+ ) >

// Strings
| <STRING:
  "\"\"
  (
    (~[\"\\\", \"\\\", \"\\n\", \"\\r\"]
    | ( \"\\\"
      ( [\"n\", \"t\", \"b\", \"r\", \"f\", \"\\\", \"'\", \"\\\" ]
      | [\"0"-"7"] ( [\"0"-"7"] )?
      | [\"0"-"3\"] [\"0"-"7\"] [\"0"-"7\"]
      )
    )
  ) *
  "\"\"
>

// Identifiers
| <ASID: <LC_LETTER> (<LETTER>|<DIGIT>|\"_\")* >
| <ID:  <LETTER>   (<LETTER>|<DIGIT>|\"_\")* >

// Path
| <PATH: ( \"/\" | \"/\" | \"\\\" | (<DRIVER> ) )
| ( (<LETTER>|<DIGIT>|\"_\")* ( \"/\" | \"\\\" ) ) * >
| <DRIVER: (<LETTER> \":\")>

| <LETTER: ( <LC_LETTER> | <UP_LETTER> ) >
| <LC_LETTER: [\"a\"-\"z\"]>
| <UP_LETTER: [\"A\"-\"Z\"]>
| <DIGIT: [\"0\"-\"9\"]>
}
```



### B.4 Source Code of Agents for Playing the Prisoner's Dilemma

This section shows the main code of the MAS configuration (Figure B.4.1), and the participant agents (Figures B.4.2, B.4.3, B.4.4, B.4.5) in the iterated PD example. Figure B.4.1 shows the MAS configuration for the two player agents: *TESTER* and *playerAff*, but the configuration for any two other player agents is similar. Agents play 100 iterations of the PD. The MAS configuration contains a third agent *arbiter*. It is in charge of moderating rounds and of informing agents about their opponent's decisions. The agent *playerAff* is an affective agent (i.e., it has a reasoning cycle like any other agent, and also it has an affective cycle that manages its affective state). The agent *playerAff*'s decisions are highly influenced by its emotional state, because its rationality level is set to its minimum value (see line 8 of Figure B.4.2). It also has a non-cooperative personality which makes its initial affective state to be 'angry'. Its decisions result in a behavior similar to the behavior of playing according to the Tit-For-Tat (TFT) strategy, but these decisions are influenced by its current affective state. The *TESTER* (Figure B.4.3) agent plays according to the TESTER strategy, where it starts testing its opponent by betraying it. Then, if its opponent's response is to betray, it apologizes by cooperating and goes on by playing a TFT strategy. Otherwise, if its opponent's response is to cooperate, it tests again the opponent by betraying it. The *FTFT* (Figure B.4.4) agent plays according to the forgiving TFT strategy, which is similar to TFT but it requires two consecutive betrayals from the opponent before betraying. The agent *betraying* (Figure B.4.5) always plays to betray.

```
1 MAS affective_iPD {
2     infrastructure: Centralised
3
4     agents:
5         arbiter;
6         TESTER;
7         playerAff
8     agentClass jason.Em.asSemantics.EmAgent
```

## APPX B. AGENTS' GRAMMAR AND EXAMPLES

---

```
9         agentArchClass jason.Em.architecture.EmAgArch
10        emEngineClass pdclasses.pd_emEngine
11        personalityClass pdclasses.pd_personality
12        affStateClass pdclasses.pd_affect;
13
14        affcategories:   grateful(<0.68:1.0>),
15                          guilty(<0.35:0.67>),
16                          angry(<0.0:0.34>);
17 }
```

Figure B.4.1: Code of the MAS configuration file for the iterated prisoner's dilemma with two players: *TESTER* and *playerAff*.

```
1  {include ("player.asl" )}
2  // Agent concerns
3  concern1 (C1):- last_move (O, OLM) & ((OLM==c&C1=2) | (OLM==b&C1=0)).
4  concern2 (C2):- my_last_move (MLM) & ((MLM==c&C2=1) | (MLM==b&C2=0)).
5  concerns__ (V):- concern1 (C1) & concern2 (C2) & V=(C1+C2)/3.
6  // Agent personality:
7  //   Cooperativeness: 0, Rat. level: 0.0 (completely emotional)
8  personality__ : { [0], 0.0 }.
9
10 /** Recording opponent's last move  */
11 @records [atomic]
12 +!take_records (T, O, M)
13 <-      -last_move (O, _);
14         +last_move (O, M).
15
16 /** Behavior */
17 // (1) When feeling grateful, cooperate
18 @pcooperate1 [atomic, affect__ (grateful)]
19 +arrested (T, O) [source (arbiter)]
20 <-      .send (arbiter, tell, play (T, c));
21         -my_last_move (_);
22         +my_last_move (c);
23         .print ("I played c (grateful) against ", O, " at time ", T).
24 // (2) When feeling guilty, cooperate
25 @pcooperate2 [atomic, affect__ (guilty)]
26 +arrested (T, O) [source (arbiter)]
27 <-      .send (arbiter, tell, play (T, c));
28         -my_last_move (_);
29         +my_last_move (c);
30         .print ("I played c (guilty) against ", O, " at time ", T).
31 // (3) When feeling angry, betray
32 @pbetray [atomic, affect__ (angry)]
```

## B.4. Source Code of Agents for Playing the Prisoner's Dilemma

---

```
33 +arrested(T,O) [source(arbiter)]
34 <- .send(arbiter, tell, play(T,b));
35 -my_last_move(_);
36 +my_last_move(b);
37 .print("I played b(angry) against ",O," at time ",T).
38 // (4) By default betray
39 @pdefault[atomic]
40 +arrested(T,O)
41 <- .send(arbiter, tell, play(T,b));
42 -my_last_move(_);
43 +my_last_move(b);
44 .print("I played b against ",O," at time ",T).
```

Figure B.4.2: Code of the agent *playerAff* in the iterated prisoner's dilemma.

```
1 {include("player.asl")}
2 /** Recording opponent's last move */
3 @records[atomic]
4 +!take_records(T,O,M)
5 <- -last_move(O,_);
6 +last_move(O,M).
7
8 /** Behavior */
9 // (1) while waiting, play 'cooperate'
10 +arrested(T,O) [source(arbiter)]
11 : last_move(O,_) & waitingOpponent
12 <- .send(arbiter, tell, play(T,c));
13 -waitingOpponent;
14 .print("I played c against ",O," at time ",T).
15 // (2) apologize and play TFT if
16 // the opponent's last move is 'betray'
17 +arrested(T,O) [source(arbiter)]
18 : last_move(O,b) & not playingTFT
19 <- .send(arbiter, tell, play(T,c));
20 +playingTFT;
21 .print("I played c against ",O," at time ",T).
22 // (3) test to betray if the
23 // opponent's last move is 'cooperate'
24 +arrested(T,O) [source(arbiter)]
25 : last_move(O,c) & not playingTFT
26 <- .send(arbiter, tell, play(T,b));
27 +waitingOpponent;
28 .print("I played b against ",O," at time ",T).
29 // (4) otherwise play TFT strategy
30 +arrested(T,O) [source(arbiter)]
31 : last_move(O,LM) & playingTFT
```

## APPX B. AGENTS' GRAMMAR AND EXAMPLES

---

```
32 <- .send(arbiter, tell, play(T,LM));
33 .print("I played ",LM," against ",O," at time ",T).
34 // (5) betray the first time
35 +arrested(T,O) [source(arbiter)]
36 <- .send(arbiter, tell, play(T,b));
37 +waitingOpponent;
38 .print("I played b against ",O," at time ",T).
```

Figure B.4.3: Code of the agent *TESTER* in the iterated prisoner's dilemma.

```
1 {include("player.asl")}
2 /** Recording opponent's last move ***/
3 // Record two last moves
4 @records1[atomic]
5 +!take_records(T,O,M)
6 : last_move2(O,M2) & last_move1(O,M1)
7 <- -last_move2(O,M2);
8 -last_move1(O,M1);
9 +last_move2(O,M1);
10 +last_move1(O,M).
11 // Record last move
12 @records2[atomic]
13 +!take_records(T,O,M)
14 : last_move1(O,M1)
15 <- -last_move1(O,M1);
16 +last_move2(O,M1);
17 +last_move1(O,M).
18 // Record first move
19 @records3[atomic]
20 +!take_records(T,O,M)
21 <- +last_move1(O,M).
22
23 /** Behavior ***/
24 // (1) betray after two betrays of the opponent
25 +arrested(T,O) [source(arbiter)]
26 : last_move2(O,b) & last_move1(O,b)
27 <- .send(arbiter, tell, play(T,b));
28 .print("I played b against ",O," at time ",T).
29 // (2) otherwise play 'cooperate'
30 +arrested(T,O) [source(arbiter)]
31 <- .send(arbiter, tell, play(T,c));
32 .print("I played c against ",O," at time ",T).
```

Figure B.4.4: Code of the agent *FTFT* in the iterated prisoner's dilemma.

## B.5. Source Code of Agents for Playing the Trust Game

---

```
1 { include("player.asl" ) }
2 // I don't need records
3 +!take_records(T,O,M) .
4
5 /** Behavior **/
6 // I always play d
7 +arrested(T,O) [source(arbiter)]
8   <- .send(arbiter, tell, play(T,b));
9     .print("I played d, as always, against ",O," at time ",T).
```

Figure B.4.5: Code of the agent *betraying* in the iterated prisoner's dilemma.

## B.5 Source Code of Agents for Playing the Trust Game

This section shows the main code of the participant agents in the iterated TG example. The MAS configuration in this example is similar to that of the PD example. The participating agents are *trustor*, *trustee*, and *moderator*, which play 6 rounds. In this example, only the agent *trustor* is affective, and the agent *trustee* sends back random amounts of endowments. It is based on the study performed in [Harth & Regner, 2016], where authors conclude that the emotion ‘anger’ and the personality trait ‘negative reciprocity’ are correlated with the difference between the back transfers and the amount sent in a previous iteration. The study also concludes that subsequent amount of endowment sent are influenced by these two factors. In the example, despite that there is a single way of behaving affectively (i.e., there is a single plan with the affective annotation), the endowments sent by the agent *trustor*'s are calculated by using the agent affective state.

## APPX B. AGENTS' GRAMMAR AND EXAMPLES

---

```
1 { include("player.asl" )
2 concerns__ (V) :- last_difference(T,Ld) &
3                 V = 0.28 * Ld - 0.24 * T .
4 personality__ : { [6,7], 0.0 }.
5
6 @records[atomic]
7 +!take_records(T,AmountSent,AmountReceived)
8   <-   -last_difference(__,__);
9         Ld = AmountReceived-AmountSent;
10        +last_difference(T,Ld) .
11
12 // By default I don't trust
13 @psend[atomic]
14 +newRound(T,O) [source(moderator)]
15   :     minimumPossible(MP) & initialEndowment(IE)
16   <-   +amountSent(T,MP);
17         Al = IE-MP;
18         +amountLeft(T,Al);
19         .send(moderator, tell, play(T,MP));
20         .print("I send ",MP," to the trustee at time ",T) .
21 // The amount I send depends on my mood
22 @psendaffective[atomic,affect__(singleCat)]
23 +newRound(T,O) [source(moderator)]
24   :     .get_affst([Afst]) & .ground(Afst) & Tant = T-1 &
25         last_difference(Tpre,Ld) & initialEndowment(IE) &
26         minimumPossible(MP)
27   <-   A = 0.24*Ld + Afst - 0.31*Tpre + 5.98;
28         if (A<0) {NewA=MP;}
29         else{ if (A>IE){NewA=IE;} else { NewA = A; } }
30         +amountSent(T,NewA);
31         Al = IE-NewA;
32         +amountLeft(T,Al);
33         .send(moderator, tell, play(T,NewA));
34         .print("I send ",NewA," to the trustee at time ",T) .
35
36 +amountReceived(T,P) [source(P)] .
```

Figure B.5.6: Code of the agent *trustor* in the iterated trust game.

## B.5. Source Code of Agents for Playing the Trust Game

---

```
1 { include("player.asl" ) }
2 // No records to take
3 +!take_records(T,AmountSent,AmountReceived) .
4
5 // The amount I send back to the trustor
6 @psendback[atomic]
7 +amountReceived(T,AmountReceived) [source(moderator)]
8 : AvailableAmount=3*AmountReceived & minimumPossible(MP)
9 <- .random(R);
10 E = R * AvailableAmount;
11 if (E<MP){ NewE = MP; } else { NewE = E; }
12 +amountSent(T,NewE);
13 A1 = AvailableAmount-NewE;
14 +amountLeft(T,A1);
15 .send(moderator, tell, play(T,NewE));
16 .print("I send ",NewE," to the trustee at time ",T).
```

Figure B.5.7: Code of the agent *betraying* in the iterated trust game.





---

---

## References

- ADAM, C. & LONGIN, D. (2007). Endowing emotional agents with coping strategies: from emotions to emotional behaviour. In *Intelligent Virtual Agents*, 348–349, Springer. 48
- ADOLPHS, R. (2003). Cognitive neuroscience of human social behaviour. *Nature Reviews Neuroscience*, **4**, 165–178. 115
- ALBEROLA, J.M., SUCH, J.M., ESPINOSA, A., BOTTI, V. & GARCÍA-FORNES, A. (2008). Magentix: a multiagent platform integrated in Linux. In *EUMAS*, 1–10. 6
- ALFONSO, B., VIVANCOS, E., BOTTI, V. & GARCÍA-FORNES, A. (2011). Integrating Jason in a multi-agent platform with support for interaction protocols. In *Proceedings of the Compilation of the Co-located Workshops on AGERE!'11, SPLASH '11 Workshop*, 221–226, ACM, NY, USA. 6, 7
- ALFONSO, B., BOTTI, V., GARRIDO, A. & GIRET, A. (2014a). A MAS-based infrastructure for negotiation and its application to a water-right market. *Information Systems Frontiers*, **16**, 183–199. 6, 113
- ALFONSO, B., VIVANCOS, E. & BOTTI, V.J. (2014b). An open architecture for affective traits in a BDI agent. In *Proceedings of the 6th International Conference on Evolutionary Computation Theory and Applications (ECTA). Part of the 6th International Joint Conference on Computational Intelligence (IJCCI)*, 320–325. 61

## REFERENCES

---

- ALFONSO, B., PYNADATH, D.V., LHOMMET, M. & MARSELLA, S. (2015a). Emotional perception for updating agents' beliefs. In *International Conference on Affective Computing and Intelligent Interaction (ACII)*, 201–207, IEEE. 180
- ALFONSO, B., VIVANCOS, E., BOTTI, V. & HERNÁNDEZ, P. (2015b). Building emotional agents for strategic decision-making. In *Proceedings of the 7th International Conference on Agents and Artificial Intelligence (ICAART)*, 390–397. 116
- ALFONSO, B., DEL VAL, E. & ALBEROLA, J.M. (2016a). An agent model of student's affect for adapting schooling strategies. In *Methodologies and Intelligent Systems for Technology Enhanced Learning, 6th International Conference*, 15–22, Springer. 5
- ALFONSO, B., VIVANCOS, E. & BOTTI, V.J. (2016b). Design of an affective intelligent agent on *GenIA*<sup>3</sup>. Tech. rep., DSIC, UPV, Spain. 103
- ALFONSO, B., VIVANCOS, E. & BOTTI, V.J. (2016c). Toward a systematic development of affective intelligent agents. Tech. rep., DSIC, UPV, Spain. 61
- ALFONSO, B., VIVANCOS, E. & BOTTI, V. (2017). Toward formal modeling of affective agents in a BDI architecture. *ACM Transactions on Internet Technology (TOIT)*, **17**, Article 5. 61, 79
- ALLPORT, G.W. (1937). *Personality: a psychological interpretation*. Henry Holt New York. 28, 89
- ALLPORT, G.W. (1961). *Pattern and growth in personality*. Outstanding books in education and psychology, Holt, Rinehart and Winston. 28
- ALLPORT, G.W., ODBERT, H.S. & LABORATORY, H.P. (1936). *Trait-names: a psycho-lexical study*. Psychological Review Publications. Psychological Monographs. v. 47, no. 1; whole no. 211, Psychological Review Company. 28
- ARNOLD, M.B. (1960). *Emotion and personality*. Columbia University Press. 23

- AVERILL, J.R. (1980). A constructivist view of emotion. In Plutchik & H. Kellerman, eds., *Emotion: Theory, Research and Experience*, vol. 1, 305–339, Academic Press, New York. 22, 24
- AXELROD, R. & HAMILTON, W.D. (1981). The evolution of cooperation. *Science*, **211**, 1390–1396. 116, 174
- AXELROD, R.M. (1984). *The evolution of cooperation*. Basic books, Basic Books. 117
- AYLETT, R. & PAIVA, A. (2012). Computational modelling of culture and affect. *Emotion Review*, **4**, 253–263. 9, 53, 77
- BANDURA, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, **84**, 191. 67
- BANDURA, A. (1996). Ontological and epistemological terrains revisited. *Journal of Behavior Therapy and Experimental Psychiatry*, **27**, 323–345. 134
- BARON, R.A., BYRNE, D.E. & BRANSCOMBE, N.R. (2006). *Social psychology*. MyPsychLab Series, Pearson/Allyn & Bacon. 28
- BATTAGLINO, C., DAMIANO, R. & LESMO, L. (2013). Emotional range in value-sensitive deliberation. In *Proceedings of AAMAS '13*, 769–776, International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC. 52
- BAUMEISTER, R.F., VOHS, K.D., DEWALL, C.N. & ZHANG, L. (2007). How emotion shapes behavior: feedback, anticipation, and reflection, rather than direct causation. *Personality and Social Psychology Review*, **11**, 167–203. 20
- BECHARA, A., DAMASIO, H. & DAMASIO, A.R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral cortex*, **10**, 295–307. 4, 63
- BECKER-ASANO, C. (2008). *WASABI: Affect Simulation for Agents with Believable Interactivity*, vol. 319. IOS Press. 41

## REFERENCES

---

- BECKER-ASANO, C. & WACHSMUTH, I. (2010). Affective computing with primary and secondary emotions in a virtual human. *Proceedings of the AAMAS '10*, **20**, 32–49. 9, 41, 51
- BERG, J., DICKHAUT, J. & MCCABE, K. (1995). Trust, reciprocity, and social history. *Games and economic behavior*, **10**, 122–142. 122
- BODEN, M., BRYSON, J., CALDWELL, D., DAUTENHAHN, K., EDWARDS, L., KEMBER, S., NEWMAN, P., PARRY, V., PEGMAN, G., RODDEN, T. *et al.* (2017). Principles of robotics: regulating robots in the real world. *Connection Science*, **29**, 124–129. 184
- BODENHAUSEN, G.V., KRAMER, G.P. & SÜSSER, K. (1994). Happiness and stereotypic thinking in social judgment. *Journal of Personality and Social Psychology*, **66**, 621. 26
- BORDINI, R.H. & HÜBNER, J.F. (2010). Semantics for the Jason variant of Agentspeak (plan failure and some internal actions). In *Proceedings of ECAI'10*, 635–640, IOS Press, Amsterdam, The Netherlands, The Netherlands. 79, 80
- BORDINI, R.H., HÜBNER, J.F. & WOOLDRIDGE, M. (2007). *Programming multi-agent systems in Agentspeak using Jason*. Wiley. 6, 83, 106, 120, 155, 177
- BOSSE, T., BROEKENS, J., DIAS, J. & VAN DER ZWAAN, J. (2014). *Emotion modeling*. Springer. 6
- BOUKRICA, H., WACHSMUTH, I., CARMINATI, M. & KNOEFERLE, P. (2013). A computational model of empathy: empirical evaluation. In *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference On*, 1–6. 51
- BOWER, G.H., GILLIGAN, S.G. & MONTEIRO, K.P. (1981). Selectivity of learning caused by affective states. *Journal of Experimental Psychology: General*, **110**, 451. 19

- BRACHT, J. & FELTOVICH, N. (2008). Efficiency in the trust game: an experimental study of precommitment. *International Journal of Game Theory*, **37**, 39–72. 123
- BRATMAN, M. (1987). *Intention, plans, and practical reason*. Harvard University Press. 3, 33, 62, 63
- BRAVE, S., NASS, C. & HUTCHINSON, K. (2005). Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent. *International journal of human-computer studies*, **62**, 161–178. 4
- BREAZEAL, C., BUCHSBAUM, D., GRAY, J., GATENBY, D. & BLUMBERG, B. (2005). Learning from and about others: towards using imitation to bootstrap the social understanding of others by robots. *Artificial Life*, **11**, 31–62. 44
- BRIGGS, S.R. (1989). The optimal level of measurement for personality constructs. In *Personality Psychology*, 246–260, Springer. 30
- BUSEMEYER, J.R., DIMPERIO, E. & JESSUP, R.K. (2007). *Integrating emotional processes into decision-making models*, 29–44. Oxford University Press. 4
- CAMERER, C.F., LOEWENSTEIN, G. & RABIN, M. (2011). *Advances in behavioral economics*. Princeton University Press. 4, 63, 68
- CANNON, W.B. (1927). The James-Lange theory of emotions: A critical examination and an alternative theory. *The American Journal of Psychology*, **39**, 106–124. 21
- CASTELFRANCHI, C., ROSIS, F.D., FALCONE, R. & PIZZUTILO, S. (1998). Personality traits and social attitudes in multiagent cooperation. *Applied Artificial Intelligence*, **12**, 649–675. 27
- CATTELL, R.B. (1979). *Personality and learning theory: the structure of personality in its environment*. Personality and Learning Theory, Springer Publishing Company, Incorporated, New York. 29

## REFERENCES

---

- CATTELL, R.B. & TREGASKIS, D. (1965). *The scientific analysis of personality*, vol. 27. Penguin Books Harmondsworth. 29
- CHARNESS, G. & RABIN, M. (2002). Understanding social preferences with simple tests. *The Quarterly Journal of Economics*, **117**, 817–869. 115
- CLARK, K. & SEFTON, M. (2001). The sequential prisoner’s dilemma: evidence on reciprocation. *The Economic Journal*, **111**, 51–68. 117
- COHEN, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, N.J. : L. Erlbaum Associates. 143
- CONWAY, M.A. (1990). *Autobiographical memory: an introduction*. Open University Press. 65
- COPPIN, G. & SANDER, D. (2016). Theoretical approaches to emotion and its measurement. In H.L. Meiselman, ed., *Emotion Measurement*, 3–30, Woodhead Publishing. 129
- COWIE, R. (2015). Ethical issues in affective computing. *The Oxford Handbook of Affective Computing*, 334. 183, 184
- DAMÁSIO, A.R. (1994). *Descartes’ error: emotion, reason, and the human brain*. Quill. 5, 18, 20, 22, 41, 45, 73
- DARWIN, C. (1872). *The expression of the emotions in man and animals*. John Murray, the original was published 1898 by Appleton, New York. Reprinted 1965 by the University of Chicago Press. 21
- DAVIES, J.D. (1955). *Phrenology, fad and science: a 19th century american crusade*. Yale University Press. 28
- DE SOUSA, R. (1990). *The rationality of emotion*. Mit Press. 77
- DE WAAL, F.B. (2008). Putting the altruism back into altruism: the evolution of empathy. *Annual Review of Psychology*, **59**, 279–300. 27

- DEMAREE, H.A., DEDONNO, M.A., BURNS, K.J., FELDMAN, P. & EVERHART, D.E. (2009). Trait dominance predicts risk-taking. *Personality and Individual Differences*, **47**, 419–422. 20
- DEMOS, R. (1960). Lying to oneself. *The Journal of Philosophy*, **57**, 588–595. 26
- DIAS, J., MASCARENHAS, S. & PAIVA, A. (2014). *FAtiMa Modular: Towards an Agent Architecture with a Generic Appraisal Framework*, chap. Generic Models and Frameworks, 44–56. Springer International Publishing, Cham. 9, 53, 69, 77, 176
- DIGMAN, J.M. & TAKEMOTO-CHOCK, N.K. (1981). Factors in the natural language of personality: re-analysis, comparison, and interpretation of six major studies. *Multivariate Behavioral Research*, **16**, 149–170. 30
- DOHMEN, T., FALK, A., HUFFMAN, D. & SUNDE, U. (2009). Homo reciprocans: survey evidence on behavioural outcomes. *The Economic Journal*, **119**, 592–612. 124
- EKMAN, P. & FRIESEN, W. (1978). *Facial action coding system: a technique for the measurement of facial movement*. Consulting Psychologists Press, Palo Alto. 129
- EKMAN, P. & OSTER, H. (1982). Review of research, 1970–1980. *Emotion in the human face*, 147–173. 46
- EKMAN, P., FRIESEN, W.V., O’SULLIVAN, M., DIACOYANNI-TARLATZIS, I., KRAUSE, R., PITCAIRN, T., SCHERER, K.R., CHAN, A., HEIDER, K., LECOMPTE, W.A., RICCI-BITTI, P.E., TOMITA, M. & TZAVARAS, A. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of Personality and Social Psychology*, **53**, 712–717. 22
- EL-NASR, M.S., YEN, J. & IOERGER, T.R. (2000). Flame—fuzzy logic adaptive model of emotions. *Autonomous Agents and Multi-agent systems*, **3**, 219–257. 52

## REFERENCES

---

- ELBANNA, S. (2006). Strategic decision-making: process perspectives. *International Journal of Management Reviews*, **8**, 1–20. 5
- ELLSBERG, D. (1961). Risk, ambiguity, and the savage axioms. *The Quarterly Journal of Economics*, 643–669. 26
- EYSENCK, H.J. (1970). *The structure of human personality*. Methuen, [3d ed.] edn. 29
- EYSENCK, H.J. (1982). *Personality, genetics, and behavior: selected papers*, chap. Development of a Theory. Centennial psychology series, Praeger. 29, 89
- EYSENCK, H.J. (1990). Biological dimensions of personality. In L. Pervin, ed., *Handbook of Personality: Theory and Research*, 244–276, New York: Guilford. 75
- FOGUÉS, R.L., ALBEROLA, J.M., SUCH, J.M., ESPINOSA, A. & GARCÍA-FORNES, A. (2010). Towards dynamic agent interaction support in open multiagent systems. In *Proceedings of the 13th International Conference of the Catalan Association for Artificial Intelligence*, vol. 220 of , 89–98, IOS Press. 7
- FORGAS, J. (1995). Mood and judgment: the affect infusion model (AIM). *Psychol Bull*, **117**, 39–66. 136, 180
- FORGAS, J.P. (1998). On being happy and mistaken: mood effects on the fundamental attribution error. *Journal of Personality and Social Psychology*, **75**, 318. 26
- FRANKFURT, H.G. (1988). *The importance of what we care about*. Cambridge University Press. 23
- FRANKLIN, S., MADL, T., D’MELLO, S. & SNAIDER, J. (2014). LIDA: a systems-level architecture for cognition, emotion, and learning. *Autonomous Mental Development, IEEE Transactions on*, **6**, 19–41. 9, 39, 40



- FREDERICK, S. (2005). Cognitive reflection and decision making. *The Journal of Economic Perspectives*, **19**, 25–42. 67
- FREUD, S. (1937). *Analysis terminable and interminable*, vol. 23 of *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, 216–237. London: Hogarth Press, James Strachey edn. 25
- FRIDLUND, A.J. (1994). *Human facial expression: an evolutionary view*. Academic Press. 22
- FRIJDA, N., MANSTEAD, A. & BEM, S. (2000). *Emotions and beliefs: how feelings influence thoughts*. Studies in Emotion and Social Interaction, Cambridge University Press. 26, 49, 67
- FRIJDA, N.H. (1987). *The emotions*. Studies in Emotion and Social Interaction, Cambridge University Press. 77
- FRIJDA, N.H. (2007). *The laws of emotion*. Lawrence Erlbaum Associates, Incorporated. 23, 67, 75, 92
- GEBHARD, P. (2005). ALMA: A layered model of affect. In *Proceedings of the 4th AAMAS*, 29–36, ACM, NY, USA. 103, 105, 159
- GERRARDS-HESSE, A., SPIES, K. & HESSE, F.W. (1994). Experimental inductions of emotional states and their effectiveness: A review. *British Journal of Psychology*, **85**, 55–78. 133
- GIRET, A., GARRIDO, A., GIMENO, J.A., BOTTI, V. & NORIEGA, P. (2011). A MAS decision support tool for water-right markets. In *The 10th International Conference on Autonomous Agents and Multiagent Systems—Volume 3*, 1305–1306, International Foundation for Autonomous Agents and Multiagent Systems. 6, 113
- GOHM, C.L. *et al.* (2003). Mood regulation and emotional intelligence: individual differences. *Journal of personality and social psychology*, **84**, 594–607.

## REFERENCES

---

- GOLDBERG, L.R. *et al.* (1990). An alternative “Description of personality”: the big-five factor structure. *Journal of Personality and Social Psychology*, **59**, 1216–1229. 30, 89
- GOLLWITZER, P.M. (1999). Implementation intentions: strong effects of simple plans. *American psychologist*, **54**, 493. 20
- GREENSPAN, P. (2004). Practical reasoning and emotion. *The Oxford handbook of rationality*, 206–21. 18, 19
- GRIFIN, P.A. (1996). *The theory of Blackjack. The complete card counter’s guide to the casino game of 21*. Huntigton Press. 130, 134, 152, 168
- GROSS, J.J. & THOMPSON, R.A. (2011). Emotion regulation: conceptual foundations. In *Handbook of Emotion Regulation*, Guilford Publications. 24, 66
- HARTH, N.S. & REGNER, T. (2016). The spiral of distrust: (Non-)cooperation in a repeated trust game is predicted by anger and individual differences in negative reciprocity orientation. *International Journal of Psychology*. 123, 195
- HATFIELD, E., CACIOPPO, J. & RAPSON, R. (1992). Primitive emotional contagion. *Emotion and Social Behavior*, 151–177. 27
- HEISE, D.R. (2007). *Expressive order: confirming sentiments in social actions*. Springer Science & Business Media. 21, 43
- HESS, U. & FISCHER, A. (2014). Emotional mimicry: why and when we mimic emotions. *Social and Personality Psychology Compass*, **8**, 45–57. 27
- HESS, U., PHILIPPOT, P. & BLAIRY, S. (1998). Facial reactions to emotional facial expressions: affect or cognition? *Cognition and Emotion*, **12**, 509–531. 26
- HOEY, J., SCHRÖDER, T. & ALHOTHALI, A. (2016). Affect control processes: intelligent affective interaction using a partially observable markov decision process. *Artificial Intelligence*, **230**, 134–172. 9, 21, 43

- HOWARD, P.J. & HOWARD, J.M. (1995). The big five quickstart: an introduction to the five-factor model of personality for human resource professionals. *Center for Applied Cognitive Studies*. 27
- HU, S., BLETTNER, D. & BETTIS, R.A. (2011). Adaptive aspirations: performance consequences of risk preferences at extremes and alternative reference groups. *Strategic Management Journal*, **32**, 1426–1436. 134
- HUDLICKA, E. (2004). Two sides of appraisal: implementing appraisal and its consequences within a cognitive architecture. In *Proceedings of the AAAI Spring Symposium 2004, Architectures for Modeling Emotion*. 9, 52, 77
- HUDLICKA, E. (2008). Modeling the mechanisms of emotion effects on cognition. In *AAAI Fall Symposium: Biologically Inspired Cognitive Architectures*, 82–86. 9, 52, 77
- HUDLICKA, E. (2014). From habits to standards: towards systematic design of emotion models and affective architectures. In *Emotion Modeling*, 3–23, Springer International Publishing. 9, 57, 64
- ITO, J., PYNADATH, D. & MARSELLA, S. (2010). Modeling self-deception within a decision-theoretic framework. *AAMAS*, **20**, 3–13. 67
- JAMES, W. (1884). What is an emotion? *Mind*, 188–205. 21, 22
- JAMES, W. (1890). *The principles of psychology*. Henry Holt New York. 19
- JOHN, O.P. (1989). Towards a taxonomy of personality descriptors. In *Personality Psychology*, 261–271, Springer. 30
- JOHN, O.P. & SRIVASTAVA, S. (1999). The big five trait taxonomy: history, measurement, and theoretical perspectives. In L.A. Pervin & O.P. John, eds., *Handbook of Personality: Theory and Research*, 102–138, Guilford Press, 2nd edn. 27, 30
- KAHNEMAN, D. & MILLER, D.T. (1986). Norm theory: comparing reality to its alternatives. *Psychological review*, **93**, 136. 76

## REFERENCES

---

- KEFALAS, P., SAKELLARIOU, I., BASAKOS, D. & STAMATOPOULOU, I. (2014). A formal approach to model emotional agents behaviour in disaster management situations. In *Hellenic Conference on Artificial Intelligence*, 237–250, Springer. 5
- KENNEDY, W.G. (2012). Modelling human behaviour in agent-based models. In *Agent-based Models of Geographical Systems*, 167–179, Springer. 68
- KHADJAVI, M. & LANGE, A. (2013). Prisoners and their dilemma. *Journal of Economic Behavior & Organization*, **92**, 163–175. 117
- KIRMAN, A., LIVET, P. & TESCHL, M. (2009). Rationality and emotions. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **365**, 215–219. 19, 26
- KLEIN, J., MOON, Y. & PICARD, R.W. (2002). This computer responds to user frustration: theory, design, and results. *Interacting with computers*, **14**, 119–140. 4
- KOVAL, P., BUTLER, E.A., HOLLENSTEIN, T., LANTEIGNE, D. & KUPPENS, P. (2014). Emotion regulation and the temporal dynamics of emotions: effects of cognitive reappraisal and expressive suppression on emotional inertia. *Cognition and Emotion*, 1–21. 67
- LARSEN, R.J. (2000). Toward a science of mood regulation. *Psychological Inquiry*, **11**, 129–141. 24
- LAZARUS, R. (1994). *Emotion and adaptation*. Oxford University Press. 23, 24
- LAZARUS, R.S. & FOLKMAN, S. (1984). *Stress, appraisal, and coping*. Springer Publishing Company. 25, 48, 66
- LEDOUX, J.E. (1992). Emotion as memory: anatomical systems underlying indelible neural traces. *The handbook of emotion and memory: Research and theory*, 269–288. 19
- LEDOUX, J.E. (1998). *The emotional brain: the mysterious underpinnings of emotional life*. A Touchstone book, Simon & Schuster. 5, 22, 41

- LERNER, J.S., SMALL, D.A. & LOEWENSTEIN, G. (2004). Heart strings and purse strings: carryover effects of emotions on economic decisions. *Psychological Science*, **15**, 337–41. 5
- LESSMANN, N., KOPP, S. & WACHSMUTH, I. (2006). Situated interaction with a virtual human perception, action, and cognition. *Situated Communication*, **166**, 287. 51
- LEVINE, L.J. & PIZARRO, D.A. (2004). Emotion and memory research: A grumpy overview. *Social Cognition*, **22**, 530–554. 19
- LOEWENSTEIN, G., WEBER, E.U., HSEE, C.K. & WELCH, N. (2001). Risk as feelings. *Psychological Bulletin*, **127**, 267–286. 5
- LONGIN, C.A.D. (2007). Endowing emotional agents with coping strategies: from emotions to emotional behaviour. Tech. rep., Institut de Recherche en Informatique de Toulouse (IRIT). 48
- MANI, I. (2014). Computational narratology. *Handbook of narratology*, 84–92. 177
- MARINIER, I.R.P., LAIRD, J.E. & LEWIS, R.L. (2009). A computational unification of cognitive behavior and emotion. *Cognitive Systems Research*, **10**, 48–69. 52
- MARSELLA, S. & GRATCH, J. (2003). Modeling coping behavior in virtual humans: don't worry, be happy. In *Proceedings of AAMAS'03*, 313–320, ACM. 67
- MARSELLA, S.C. & GRATCH, J. (2009). EMA: A process model of appraisal dynamics. *Cognitive Systems Research*, **10**, 70–90. 25, 52, 65, 66, 69, 103, 105
- MARSELLA, S.C., GRATCH, J. & PETTA, P. (2010). Computational models of emotion. In *a Blueprint for Affective Computing: A Sourcebook and Manual*, Affective Science, 21–46, OUP Oxford. 52

## REFERENCES

---

- MCCRAE, R.R. & COSTA, P.T. (1997). Personality trait structure as a human universal. *American psychologist*, **52**, 509–516. 29
- MCCRAE, R.R. & JOHN, O.P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, **60**, 175–215. 29, 30, 75, 85, 90, 103, 152
- MEHRABIAN, A. (1996a). Analysis of the big-five personality factors in terms of the PAD temperament model. *Australian Journal of Psychology*, **48**, 86–92. 104
- MEHRABIAN, A. (1996b). Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. *Current Psychology*, **14**, 261–292. 41, 43, 92, 103, 152
- MEHRABIAN, A. (1997). Comparison of the PAD and PANAS as models for describing emotions and for differentiating anxiety from depression. *Journal of Psychopathology and Behavioral Assessment*, **19**, 331–357. 137, 152
- MEHRABIAN, A. & RUSSELL, J.A. (1974). *An approach to environmental psychology*. MIT Press. 66
- MEYER, J.J.C. (2006). Reasoning about emotional agents. *International Journal of Intelligent Systems*, **21**, 601–619. 47
- MEYER, J.J.C., VAN DER HOEK, W. & VAN LINDER, B. (1999). A logical approach to the dynamics of commitments. *Artificial Intelligence*, **113**, 1–40. 47
- MOORS, A. & SCHERER, K.R. (2013). The role of appraisal in emotion. *Handbook of cognition and emotion*, 135–155. 22
- MOORS, A., ELLSWORTH, P.C., SCHERER, K.R. & FRIJDA, N.H. (2013). Appraisal theories of emotion: state of the art and future development. *Emotion Review*, **5**, 119–124. 23
- MOREWEDGE, C.K. & GIBLIN, C.E. (2015). Explanations of the endowment effect: an integrative review. *Trends in cognitive sciences*, **19**, 339–348. 5

- MORRIS, W.N. (2012). *Mood: the frame of mind*. Springer Science & Business Media. 4
- MYERS, I.B., MCCAULLEY, M.H. & MOST, R. (1985). *Manual: a guide to the development and use of the myers-briggs type indicator*. Consulting Psychologists Press Palo Alto, CA. 75
- NASH, J. (1951). Non-cooperative games. *Annals of Mathematics*, 286–295. 115
- NELSON, K. (1993). The psychological and social origins of autobiographical memory. *Psychological science*, 4, 7–14. 73, 91
- NORMAN, W.T. (1967). *2800 personality trait descriptors: normative operating characteristics for a university population*. University of Michigan, Department of Psychology. 30
- OLSSON, A. & OCHSNER, K.N. (2008). The role of social cognition in emotion. *Trends in cognitive sciences*, 12, 65–71. 115
- ORTMANN, A. & TICHY, L.K. (1999). Gender differences in the laboratory: evidence from prisoner’s dilemma games. *Journal of Economic Behavior & Organization*, 39, 327–339. 117
- ORTONY, A. (2003). On making believable emotional agents believable. In R.P. Trappell, P. Petta & S. Payer, eds., *Emotions in Humans and Artifacts*, chap. 6, 189–212, MIT Press. 66, 77
- ORTONY, A., CLORE, G.L. & COLLINS, A. (1988). *The cognitive structure of emotions*. Cambridge University Press. 4, 23, 37, 41, 47, 65, 69, 72, 75, 76, 80
- ORTONY, A., NORMAN, D.A. & REVELLE, W. (2005). Affect and proto-affect in effective functioning. In *Who Needs Emotions. the Brain Meets the Machine.*, 173–302, J.M. Fellous & M.A. Arbib. 41
- PENA, L., OSSOWSKI, S., PEÑA, J.M. & SÁNCHEZ, J.A. (2011). EEP—a lightweight emotional model: application to rpg video game characters. In

## REFERENCES

---

- Computational Intelligence and Games (CIG), 2011 IEEE Conference On*, 142–149, IEEE. 5
- PICARD, R.W. (1997). *Affective computing*. MIT Press, Cambridge, MA, USA. 8
- PICARD, R.W. (2002). *What does it mean for a computer to "have" emotions?*, 213–236. MIT Press. 7
- PICARD, R.W. & LIU, K.K. (2007). Relative subjective count and assessment of interruptive technologies applied to mobile monitoring of stress. *International Journal of Human-Computer Studies*, **65**, 361–375. 4
- PIMENTEL, C. & CRAVO, M. (2005). Affective revision. In *Progress in Artificial Intelligence*, vol. 3808 of *LNCS*, 115–126, Springer Berlin Heidelberg. 26, 49, 67
- PLOTKIN, G.D. (1981). A structural approach to operational semantics. Tech. Rep. DAIMI FN-19, Aarhus University. 95
- RAO, A.S. (1996). AgentSpeak(L): BDI agents speak out in a logical computable language. In R. Van Hoe, ed., *Seventh European Workshop on Modelling Autonomous Agents in a Multi-Agent World*, Eindhoven, The Netherlands. 80
- READ, S.J., MONROE, B.M., BROWNSTEIN, A.L., YANG, Y., CHOPRA, G. & MILLER, L.C. (2010). A neural network model of the structure and dynamics of human personality. *Psychological Review*, **117**, 61. 28
- REISENZEIN, R., HUDLICKA, E., DASTANI, M., GRATCH, J., HINDRIKS, K., LORINI, E. & MEYER, J.J. (2013). Computational modeling of emotion: toward improving the inter- and intradisciplinary exchange. *IEEE Transactions on Affective Computing*, **4**, 246–266. 9, 10, 31, 34, 52, 61, 62, 75, 176
- RIZZO, P., VELOSO, M., MICELI, M. & CESTA, A. (1997). Personality-driven social behaviors in believable agents. In *Proceedings of the AAAI Fall Symposium on Socially Intelligent Agents*, 109–114. 27



- RODRÍGUEZ, L.F. & RAMOS, F. (2014). Development of computational models of emotions for autonomous agents: A review. *Cognitive Computation*, **6**, 351–375. 52
- ROGERS, C.R., KIRSCHENBAUM, H. & HENDERSON, V. (1989). *The Carl Rogers reader*, chap. A theory of therapy, personality, and interpersonal relationships, as developed in the client-centered framework. Houghton Mifflin Harcourt. 25
- ROSEMAN, I.J. (1996). Appraisal determinants of emotions: constructing a more accurate and comprehensive theory. *Cognition & Emotion*, **10**, 241–278. 24
- RUSSELL, J.A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin*, **115**, 102–141. 22
- RUSSELL, J.A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, **110**, 145–172. 4, 66, 80
- RUSSELL, J.A. & MEHRABIAN, A. (1977). Evidence for a three-factor theory of emotions. *Journal of Research in Personality*, **11**, 273–294. 41
- RYCKMAN, R.M. (2007). *Theories of personality*. PSY 235 Theories of Personality Series, Thomson/Wadsworth. 4, 25, 30
- SANDÍN, B., CHOROT, P., LOSTAO, L., JOINER, T.E., SANTED, M.A. & VALIENTE, R.M. (1999). Escalas panas de afecto positivo y negativo: validación factorial y convergencia transcultural. *Psicothema*, **11**, 37–51. 136, 137
- SANTOS, R., MARREIROS, G., RAMOS, C., NEVES, J. & BULAS-CRUZ, J. (2009). Personality, emotion and mood simulation in decision making. In *New Trends in Artificial Intelligence. 14th EPIA*. 27
- SANTOS, R., MARREIROS, G., RAMOS, C., NEVES, J. & BULAS-CRUZ, J. (2011). Personality, emotion, and mood in agent-based group decision making. *Intelligent Systems, IEEE*, **26**, 58–66. 74

## REFERENCES

---

- SCHERER, K.R. (1987). Toward a dynamic theory of emotion: the component process model of affective states. *Geneva Studies in Emotion and Communication*, **1**, 1–98. 23
- SCHERER, K.R. (2001). Appraisal considered as a process of multilevel sequential checking. *Appraisal processes in emotion: Theory, methods, research*, **92**, 120. 23, 42, 65, 69, 72
- SCHERER, K.R. (2004). Feelings integrate the central representation of appraisal-driven response organization in emotion. In *Feelings and Emotions: The Amsterdam Symposium*, 136–157. 23
- SCHERER, K.R. (2005). What are emotions? and how can they be measured? *Social Science Information*, **44**, 695–729. 23
- SCHERER, K.R., SCHORR, A. & JOHNSTON, T., eds. (2001). *A model of appraisal in the emotion system: integrating theory, research, and applications*, 68–91. Oxford University Press. 24, 65, 72
- SCHUTZ, M. (2012). The affect dilemma for artificial agents: should we develop affective artificial agents? *IEEE Transactions on Affective Computing*, **3**, 424–433. 183
- SCHWARTZ, B., WARD, A., MONTEROSSO, J., LYUBOMIRSKY, S., WHITE, K. & LEHMAN, D.R. (2002). Maximizing versus satisficing: happiness is a matter of choice. *Journal of personality and social psychology*, **83**, 1178. 137
- SCHWARZ, N. (2000). Emotion, cognition, and decision making. *Cognition & Emotion*, **14**, 433–440. 65
- SCHWENK, C.R. (1984). Cognitive simplification processes in strategic decision-making. *Strategic management journal*, **5**, 111–128. 5
- SELIMBEGOVIĆ, L., RÉGNER, I., HUGUET, P. & CHATARD, A. (2015). On the power of autobiographical memories: from threat and challenge appraisals to actual behaviour. *Memory*, 1–8. 65

- SEO, M.G., GOLDFARB, B. & BARRETT, L.F. (2010). Affect and the framing effect within individuals over time: risk taking in a dynamic investment simulation. *Academy of Management Journal*, **53**, 411–431. 5
- SEWELL, M. (2010). Emotions help solve the prisoner’s dilemma. In *Behavioural Finance Working Group Conference: Fairness, Trust and Emotions in Finance, London*, 1–2. 116, 117, 118
- SHAVER, P., SCHWARTZ, J., KIRSON, D. & O’CONNOR, C. (1987). Emotion knowledge: further exploration of a prototype approach. *Journal of Personality and Social Psychology*, **52**, 1061–1086. 22
- SHAVER, P.R., WU, S. & SCHWARTZ, J.C. (1992). Cross-cultural similarities and differences in emotion and its representation. A prototype approach. *Review of Personality and Social Psychology*, 175–212. 22
- SHELDON, W.H., DUPERTUIS, C.W. & MCDERMOTT, E. (1954). *Atlas of men, a guide for somatotyping the adult male at all ages*. Harper & Row, New York. 28
- SMITH, C.A. & LAZARUS, R.S. (1990). Emotion and adaptation. In L.A. Pervin, ed., *Handbook of Personality: Theory and Research*, 609–637. 24, 25, 65, 72, 75
- STEUNEBRINK, B.R., DASTANI, M. & MEYER, J.J.C. (2012). A formal model of emotion triggers: an approach for BDI agents. *Synthese*, **185**, 83–129. 9, 47, 76
- STRACK, F., MARTIN, L.L. & STEPPER, S. (1988). Inhibiting and facilitating conditions of the human smile: a nonobtrusive test of the facial feedback hypothesis. *Journal of Personality and Social Psychology*, **54**, 768. 46
- SWOPE, K.J., CADIGAN, J., SCHMITT, P.M. & SHUPP, R. (2008). Personality preferences in laboratory economics experiments. *The Journal of Socio-Economics*, **37**, 998–1009. 115
- TALARICO, J.M. & RUBIN, D.C. (2003). Confidence, not consistency, characterizes flashbulb memories. *Psychological Science*, **14**, 455–461. 19

## REFERENCES

---

- TAVERNER, J., ALFONSO, B., VIVANCOS, E. & BOTTI, V. (2016). Integrating expectations into Jason for appraisal in emotion modeling. In *Proceedings of the 8th International Conference on Evolutionary Computation Theory and Applications (ECTA). Part of the 8th International Joint Conference on Computational Intelligence (IJCCI)*, vol. 1, 231–238. 104, 179
- THAYER, R.E., NEWMAN, J.R. & MCCLAIN, T.M. (1994). Self-regulation of mood: strategies for changing a bad mood, raising energy, and reducing tension. *Journal of Personality and Social Psychology*, **67**, 910. 24
- TUCKER, A. (1983). The mathematics of tucker: A sampler. *The Two-Year College Mathematics Journal*, **14**, 228–232. 116
- TUPES, E.C. & CHRISTAL, R.E. (1992). Recurrent personality factors based on trait ratings. *Journal of Personality*, **60**, 225–251. 30
- TURKLE, S. (2010). In good company? on the threshold of robotic companions. *Close engagements with artificial companions: key social, psychological, ethical and design issues. Amsterdam, The Netherlands: John Benjamins Publishing Company*, 3–10. 184
- VAN KNIPPENBERG, D., KOOIJ-DE BODE, H.J. & VAN GINKEL, W.P. (2010). The interactive effects of mood and trait negative affect in group decision making. *Organization Science*, **21**, 731–744. 137
- VERDUYN, P., VAN MECHELEN, I., TUERLINCKX, F. & SCHERER, K. (2013). The relation between appraised mismatch and the duration of negative emotions: evidence for universality. *European Journal of Personality*, **27**, 481–494. 67
- VIEIRA, R., MOREIRA, Á.F., WOOLDRIDGE, M. & BORDINI, R.H. (2007). On the formal semantics of speech-act based communication in an agent-oriented programming language. *Journal of Artificial Intelligent Research (JAIR)*, **29**, 221–267. 13, 79, 80, 81, 92, 94, 95, 97, 100

- WATSON, D., CLARK, L.A. & TELLEGEN, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology*, **54**, 1063. 136
- WEISS, G. (1999). *Multiagent systems: a modern approach to distributed artificial intelligence*. MIT press. 33, 34
- WEISS, G. (2013). *Multiagent systems*. Intelligent robotics and autonomous agents, MIT Press. 63
- WHITEN, A. (1991). *Natural theories of mind: evolution, development and simulation of everyday mindreading*. Basil Blackwell Oxford. 26
- WILLIAMS, A.C.D.C. (2002). Facial expression of pain, empathy, evolution, and social learning. *Behavioral and Brain Sciences*, **25**, 475–480. 26
- ZANON, C., BASTIANELLO, M.R., PACICO, J.C. & HUTZ, C.S. (2013). Relationships between positive and negative affect and the five factors of personality in a brazilian sample. *Paidéia (Ribeirão Preto)*, **23**, 285–292. 152
- ZELENSKI, J.M. (2007). The role of personality in emotion, judgment and decision making. *Vohs K, Baumeister R, Loewenstein G, Do Emotions Help or Hurt Decision Making*. 7



---

## Subject Index

- affect, 4
- affective category, 80
- affective characteristics, 3
- affective computing, 8
- affective cycle, 80
- affective processes, 80
- affective state, 4, 80
- agents, 3
- appraisal, 22
- appraisal variables, 23
- BDI, 3
- concerns, 23
- coping strategies, 24
- core affect, 4, 80
- deliberation, 34
- emotion category, 80
- emotion regulation, 24
- emotions, 4
- MAS, 3
- means-end reasoning, 34
- mood, 4, 41
- personality, 4, 27
- personality traits, 27
- practical reasoning, 3, 33
- primary emotions, 41
- reasoning cycle, 80
- secondary emotions, 41
- theory of mind, 26