Autonomous Decision-Making based on Biological Adaptive Processes for Intelligent Social Robots

by

Marcos Maroto Gómez

A dissertation submitted by in partial fulfilment of the requirements for the degree of Doctor of Philosophy in
Electrical Engineering, Electronics and Automation

Universidad Carlos III de Madrid

Advisors:

María Malfaz Vázquez
Álvaro Castro González

Tutor:

Miguel Ángel Salichs Sánchez-Caballero

March 2022

To my loved family and those who always believed in me.

Acknowledgements

After more than four years of unique experiences combined with a roller-coaster of feelings, I would like to express my gratitude to the beautiful people that had been by my side in good and bad times. I take the liberty of writing the following lines in Spanish, so everybody to whom they are directed will understand how happy I feel by expressing my gratitude to them.

En primer lugar, me gustaría agradecer a toda mi familia su apoyo en todos los momentos importantes de mi vida. Especialmente, me gustaría destacar a mis padres, José y Elena, y a mi hermano Jorge, ya que sin su apoyo y esfuerzo nunca podría haber logrado todo lo que he conseguido ni en el ámbito personal ni en el académico. Continúo destancando el papel fundamental de mis abuelos, que durante los últimos años tan orgullos se han mostrado de lo que he conseguido allá donde iban. Pese a que no todos estaréis en este logro tan importante para mí, espero que podáis ver allí donde estéis todo lo que sigo consiguiendo gracias a vosotros. No me olvido de mis primos Ángela, Clau, Andrea, Lucía y Hugo, ya que son una parte fundamental de mi vida. Tampoco de mis tios, especialmente de Prados, por todo lo que han hecho por mí. Por último, destacar el papel de Laura por todos los momentos que hemos pasado juntos y la paciencia que ha tenido durante todos ellos. Sin todos vosotros, nada de esto hubiese sido posible.

Me gustaría dar las gracias a mis directores de tesis, María y Álvaro, y a mi tutor Miguel Ángel, por darme esta estupenda oportunidad de crecer académicamente y como persona. También me gustaria agradecer a Fernando y José Carlos por estar siempre disponibles para ayudarme en todo lo que he necesitado. No tengo palabras para agradeceros a todos vuestros consejos y el tiempo que habéis dedicado para que la realización de esta tesis haya sido posible.

Continúo agradeciendo a todos mis compañeros del grupo de robótica social con los que tanto tiempo he compartido, y tan buenos ratos he pasado, Jony, Pelayo, Juan, Sergio, Elena, Carlos, Sara Carrasco, Jaime, Abel, Quique, Juanjo y Sara Marqués. Además, me gustaría agradecer especialmente a Carlos Manuel por ser un gran amigo y apoyarme en todos los largos ratos que hemos compartido durante los últimos tres años. Incluyo en este agradecimiento a mis compañeros del departamento de Ingeniería de Sistemas y Automática, Ángela, Fernando, Raúl Nouredine, Raúl Fernández, Pavel, Jesús, Fran, Abdullah, Raúl Santos, Sonia y Edu.

Finalmente, no quiero dejar de mencionar a todos mis amigos de Toledo por los buenos momentos compartidos. En los últimos años no nos hemos visto tanto como me hubiese gustado, pero no me olvido de ninguno de vosotros, Abel, Pedro, Héctor, Darío, Jesús, Borja, Dani, Alejandro, Jony, Alberto y Enrique.

I come back to English to remember my colleagues at the Emotion, Cognition, and Inter-Action Lab at the University of Hertfordshire in London. Thank you very much to Imy and Matt for the great time we spent together and your advice. I want to thank my supervisor Lola Cañamero for the excellent opportunity to work with her and learn so many exciting things. I would not like to forget my housemates Adrien, Fauzia, and Sayo, for your support and friendship during my stay in England; you will always have a place in my heart.

¡Muchísimas gracias a todos de corazón!

Published and submitted work

Published works

Journal:

1. Maroto-Gómez, M., Castro-González, Á., Castillo, J. C., Malfaz, M., & Salichs, M. A. (2018). A bio-inspired motivational decision making system for social robots based on the perception of the user. Sensors, 18(8), 2691. Section 10.1 includes a detailed description of this work.

2. Fernández-Rodicio, E., Castro-González, Á., Alonso-Martín, F., Maroto-Gómez, M., & Salichs, M. Á. (2020). Modelling multimodal dialogues for social robots using communicative acts. Sensors, 20(12), 3440.

3. Salichs, M. A., Castro-González, Á., Salichs, E., Fernández-Rodicio, E., Maroto-Gómez, M., Gamboa-Montero, J. J., ... & Malfaz, M. (2020). Mini: a new social robot for the elderly. International Journal of Social Robotics, 12(6), 1231-1249.

4. Maroto-Gómez, M., González, R., Castro-González, Á., Malfaz, M., & Salichs, M. Á. (2021). Speeding-up Action Learning in a Social Robot with Dyna-Q+: A Bioinspired Probabilistic Model Approach. IEEE Access. Section 10.2 includes this functionality in our system.

5. Maroto-Gómez, M., Castro-González, Á., Castillo, J. C., Malfaz, M. & Salichs, M. Á. (2022). An Adaptive Decision-Making System Supported on User Preference Predictions for Human-Robot Interactive Communication. User Modeling and User-Adapted Interaction. This works is described in Section 9.2.

Conference:

1. Gómez, C. M., Juan de Dios Ursúa, C., Castillo Montoya, J. C., Castro González, Á., Alonso Martín, F., Malfaz, M., ... & Salichs, M. (2019). Desarrollo de una versión de bajo coste del robot social Mini. In XL Jornadas de Automática (pp. 718-725). Universidade da Coruña, Servizo de Publicacións.

2. Maroto, M., Castillo, J. C., Alonso Martín, F., Gamboa, J. J., Marqués Villaroya, S., & Salichs, M. Á. (2017). Definiendo los elementos que constituyen un robot social portable de bajo coste. Actas de las XXXVIII Jornadas de Automática.

3. Salichs, M., Alonso Martín, F., Malfaz, M., Castillo, J., Salichs, E., Castro-González, Á., ... & Rodicio, E. F. (2017). Interacción humano robot en el proyecto ROBSEN. presented in Jornadas Nacionales de Robótica, Valencia, 8-9.

4. Gamboa, J. J., Alonso Martín, F., Castillo, J. C., Marqués Villaroya, S., Maroto, M., & Salichs, M. Á. (2017). Micrófonos de contacto: una alternativa para sensado tactil en robots sociales. Actas de las XXXVIII Jornadas de Automática.

5. Marqués Villaroya, S., Castillo, J. C., Alonso Martín, F., Maroto, M., Gamboa, J. J., & Salichs, M. Á. (2017). Interfaces táctiles para Interacción Humano-Robot. In XXXVIII Jornadas de Automática (pp. 787-792). Servicio de Publicaciones de la Universidad de Oviedo.

6. Alonso Martín, F., Castillo, J. C., Castro-González, Á., Gamboa, J. J., Maroto, M., Marqués Villaroya, S., ... & Salichs, M. Á. (2017). Percibiendo el entorno en los robots sociales del RoboticsLab. Actas de las XXXVIII Jornadas de Automática.

Submitted works

Journal:

1. Maroto-Gómez, M., Lewis, M., Castro González, Á., Malfaz, M., Salichs, M. Á. & Cañamero, L. (2021). Adapting to My User, Engaging with My Robot: An Adaptive Affective Architecture for an Interactive Assistive Social Robot. International Journal of Human Computer Studied. This work is described in Chapter 8.

2. Fernández-Rodicio, E., Maroto-Gómez, M., Castro González, Á., Malfaz, M. & Salichs, M. Á. (2021). Emotion and Mood Regulation in Embodied Artificial Agents: Expressing Affective States in the social robot Mini. International Journal of Social Robotics. This work is fully described in Section 12.1.

3. Maroto-Gómez, M., Castro-González, Á., Malfaz, M., & Salichs, M. Á. (2021). A Motivational Model based on Artificial Biological Functions for the Intelligent Decision-Making of Social Robots. Memetic Computing. The model presented in this paper can be found in Chapter 6.

4. Maroto-Gómez, M., Alonso Martín, F., Malfaz, M., Castro González, Á., Castillo, J. C. & Salichs, M. Á. (2022). A Literature Review of Decision-Making and Control Systems for Autonomous and Social Robots. International Journal of Social Robotics. This works is described in Section 3.

Other research merits

- Visiting researcher at the Embodied Emotion, Cognition, and Inter-Action Lab, University of Hertfordshire, England. From April 2019 to August 2019.
- Local Organising Committee in 11th International Conference on Social Robotics (ICSR), Madrid, Spain.
- Volunteering in the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Madrid, Spain.
- Four-year (2017-2021) internship (PIPF) for coursing PhD studies at Department of Systems Engineering and Automation, University Carlos III of Madrid.

Abstract

The unceasing development of autonomous robots in many different scenarios drives a new revolution to improve our quality of life. Recent advances in human-robot interaction and machine learning extend robots to social scenarios, where these systems pretend to assist humans in diverse tasks. Thus, social robots are nowadays becoming real in many applications like education, healthcare, entertainment, or assistance. Complex environments demand that social robots present adaptive mechanisms to overcome different situations and successfully execute their tasks. Thus, considering the previous ideas, making autonomous and appropriate decisions is essential to exhibit reasonable behaviour and operate well in dynamic scenarios.

Decision-making systems provide artificial agents with the capacity of making decisions about how to behave depending on input information from the environment. In the last decades, human decision-making has served researchers as an inspiration to endow robots with similar deliberation. Especially in social robotics, where people expect to interact with machines with human-like capabilities, biologically inspired decision-making systems have demonstrated great potential and interest. Thereby, it is expected that these systems will continue providing a solid biological background and improve the naturalness of the human-robot interaction, usability, and the acceptance of social robots in the following years.

This thesis presents a decision-making system for social robots acting in healthcare, entertainment, and assistance with autonomous behaviour. The system's goal is to provide robots with natural and fluid human-robot interaction during the realisation of their tasks. The decision-making system integrates into an already existing software architecture with different modules that manage human-robot interaction, perception, or expressiveness. Inside this architecture, the decision-making system decides which behaviour the robot has to execute after evaluating information received from different modules in the architecture. These modules provide structured data about planned activities, perceptions, and artificial biological processes that evolve with time that are the basis for natural behaviour. The natural behaviour of the robot comes from the evolution of biological variables that emulate biological processes occurring in humans. We also propose a Motivational model, a module that emulates biological processes in humans for generating an artificial physiological and psychological state that influences the robot's decision-making. These processes emulate the natural biological rhythms of the human

organism to produce biologically inspired decisions that improve the naturalness exhibited by the robot during human-robot interactions. The robot's decisions also depend on what the robot perceives from the environment, planned events listed in the robot's agenda, and the unique features of the user interacting with the robot.

The robot's decisions depend on many internal and external factors that influence how the robot behaves. Users are the most critical stimuli the robot perceives since they are the cornerstone of interaction. Social robots have to focus on assisting people in their daily tasks, considering that each person has different features and preferences. Thus, a robot devised for social interaction has to adapt its decisions to people that aim at interacting with it. The first step towards adapting to different users is identifying the user it interacts with. Then, it has to gather as much information as possible and personalise the interaction. The information about each user has to be actively updated if necessary since outdated information may lead the user to refuse the robot. Considering these facts, this work tackles the user adaptation in three different ways.

- The robot incorporates user profiling methods to continuously gather information from the user using direct and indirect feedback methods.
- The robot has a Preference Learning System that predicts and adjusts the user's preferences to the robot's activities during the interaction.
- An Action-based Learning System grounded on Reinforcement Learning is introduced as the origin of motivated behaviour.

The functionalities mentioned above define the inputs received by the decision-making system for adapting its behaviour. Our decision-making system has been designed for being integrated into different robotic platforms due to its flexibility and modularity. Finally, we carried out several experiments to evaluate the architecture's functionalities during real human-robot interaction scenarios. In these experiments, we assessed:

- How to endow social robots with adaptive affective mechanisms to overcome interaction limitations.
- Active user profiling using face recognition and human-robot interaction.
- A Preference Learning System we designed to predict and adapt the user preferences towards the robot's entertainment activities for adapting the interaction.
- A Behaviour-based Reinforcement Learning System that allows the robot to learn the effects of its actions to behave appropriately in each situation.
- The biologically inspired robot behaviour using emulated biological processes and how the robot creates social bonds with each user.

- The robot's expressiveness in affect (emotion and mood) and autonomic functions such as heart rate or blinking frequency.

Keywords: Decision-making system; Social robots; Autonomy; Human-Robot Interaction; Intelligence; Biologically inspired models.

Contents

Acknowledgements **v**

Published and submitted content **vii**

Other research merits **ix**

Abstract **xi**

List of Figures **xxii**

List of Tables **xxxi**

Acronyms **xxxiii**

I Introduction **1**

1 Introduction **5**

1.1. An introduction to decision-making 5
1.2. Context 6
1.3. Motivation 8
1.4. The problem 9
1.4.1. Generation of natural behaviours in social robots 10
1.4.2. Robot adaptation to users 11
1.4.3. User acceptance and robot usability 12
1.5. Objectives 12
1.6. Content overview 14

2 Decision-making in humans: biological foundations 17
2.1. The nervous system 17
2.2. Homeostasis and allostasis 19
2.3. Physiological and psychological aspects of human behaviour 21
2.3.1. Physiology 21
2.3.2. Psychology 27

3 Decision-making in autonomous social robots 33
3.1. Autonomous social robots 33
3.1.1. Social robots for research 33
3.1.2. Commercial social robots for a general audience 40
3.2. Biologically inspired decision-making architectures for social robots 48
3.3. Our contribution 57

4 Learning and adaptation in social robots 59
4.1. Learning and adaptation is social robotics 59
4.2. Our approach 62
4.3. Learning methods used in this contribution 63
4.3.1. Reinforcement Learning 63
4.3.2. Preference learning. 66

II The proposed system 69

5 Our social robots 73
5.1. Our social robots 73
5.1.1. Mini 73
5.1.2. Huggable robot 75
5.2. Software architecture 76
5.2.1. The multimodal HRI system 78
5.2.2. The User-adaptive system 83
5.2.3. Skills 85

6 Our motivational model 89

6.1. Modelling neuroendocrine responses and their interactions 90
6.2. Modelling the influence of stimuli 92
6.3. Modelling circadian rhythms 95
6.4. Modelling ultradian rhythms 97
6.5. Modelling somatic and autonomic processes 98
6.6. Modelling motivation 101
6.7. Mini's behaviour 102
6.8. Modulating the robot's expressiveness with emotion and mood 103
6.9. Conclusion 104

7 Our decision-making system 107

7.1. General overview. 107
7.2. The operation of the Decision-making System 109
7.3. Conceptual design 112
7.3.1. Central controller. 113
7.3.2. Managers 114
7.3.3. Skills 115
7.4. Implementation. 117
7.4.1. Communication between the Central controller and the Managers 117
7.4.2. Communication between a Manager and a Skill 119
7.4.3. Communications with the external modules with influence on decision-making. 121
7.5. Conclusion 130

III Experiments and results 131

8 Evaluating the robot's affective adaptation 135

8.1. The social robot MiRo. 135
8.2. Robot architecture 136
8.3. Motivational model 137
8.4. Robot behaviours. 140

8.5. Emotional model 141
8.5.1. Valence dimension 141
8.5.2. Arousal dimension 143
8.5.3. Robot affect 144
8.6. MiRo's decision-making 145
8.7. Expressing affective behaviours 147
8.8. Aim of the affective adaptive mechanism 148
8.9. Experimental setup and evaluation 149
8.9.1. Scenario 150
8.9.2. User profiles 150
8.9.3. Robot profiles 151
8.10. Results 152
8.10.1. Non-adaptive robot in simulated domains 153
8.10.2. Adaptive robot in simulated domains 156
8.10.3. Non-adaptive robot in real HRIs 158
8.10.4. Adaptive robot in real human-robot interaction 161
8.11. Discussion 163

9 Evaluating the robot's Preference Learning System 165

9.1. User profiling 165
9.1.1. Profiling method 166
9.1.2. Experimental setup 168
9.1.3. Evaluation 169
9.1.4. Results 170
9.1.5. Discussion 172
9.2. Estimating the user's preferences 173
9.2.1. Designing the learning environment 173
9.2.2. Experimental setup 175
9.2.3. Evaluation 176
9.2.4. Results 178
9.2.5. Discussion 184

9.3. Adjusting the initial estimations . . . 185
9.3.1. Defining the learning scenario . . . 185
9.3.2. Experimental setup . . . 185
9.3.3. Evaluation . . . 188
9.3.4. Results . . . 190
9.3.5. Discussion . . . 194
9.4. Conclusion . . . 195

10 Evaluating the robot's Behaviour-based Learning System 197
10.1. Q-learning scenario . . . 197
10.1.1. Aim overview . . . 197
10.1.2. Defining the Q-learning environment . . . 198
10.1.3. Experimental setup . . . 201
10.1.4. Evaluation . . . 202
10.1.5. Results . . . 203
10.2. Dyna-Q+ scenario . . . 204
10.2.1. Aim overview . . . 204
10.2.2. Defining the Dyna-Q+ environment . . . 205
10.2.3. Experimental setup and evaluation . . . 208
10.2.4. Results . . . 210
10.3. Conclusion . . . 214

11 Evaluating the robot's biologically inspired behaviour 217
11.1. Motivated behaviour . . . 217
11.1.1. Mini's artificial biological processes . . . 218
11.1.2. Experimental setup and evaluation . . . 219
11.1.3. Results . . . 221
11.1.4. Discussion . . . 231
11.2. Human-robot pair-bonding . . . 233
11.2.1. Pair-bonding modelling . . . 233
11.2.2. Experimental setup and evaluation . . . 234
11.2.3. Results . . . 236

11.2.4. Conclusion 242

12 Evaluating the robot's expressiveness 243

12.1. Generating and expressing affective states 243

12.1.1. Generating affective states 244

12.1.2. Experimental setup and evaluation 245

12.1.3. Results 249

12.1.4. Discussion 253

12.2. Generating and expressing autonomic processes. 254

12.2.1. Generating autonomic processes 255

12.2.2. Experimental setup 255

12.2.3. Evaluation 256

12.2.4. Results 256

12.2.5. Discussion 267

12.3. Conclusion 267

IV Conclusion 269

13 Discussion and limitations 273

13.1. General discussion 273

13.2. Limitations 279

14 Conclusions and future work 283

14.1. Conclusions 283

14.2. Future work 285

Bibliography 287

Appendices 315

A Evaluation metrics 317

A.1. Kendall τ-b 317

A.2. Spearman ρ . . . 317
A.3. Gini impurity . . . 318
A.4. Information gain . . . 318
A.5. Mean Squared Error (MSE) . . . 318

B Methodologies **319**
B.1. Boltzmann's action selection method . . . 319
B.2. K-fold cross-validation . . . 319
B.3. Hyperparameter optimization methods . . . 320

C UML diagrams **321**

D Questionnaires **323**
D.1. User profiling questions . . . 323
D.2. Ad-hoc questions . . . 325
D.3. Robotic Social Attributes Scale (RoSAS) . . . 326
D.4. System Usability Scale (SUS) . . . 327
D.5. Godspeed series . . . 327

List of Figures

2.1 Structure of the human nervous system 18

2.2 Parasympathetic and sympathetic functions 19

2.3 Homeostatic and allostatic regulatory mechanisms 20

2.4 The melatonergic and orexinergic systems in the human brain 24

2.5 The limbic system as the main emotion regulator in the human brain . . . 25

2.6 Brain localization of oxytocin and arginine vasopressin 26

2.7 The human HPA axis 26

2.8 Lövheim's three dimensional model of affect implmented in NEUCOGAR 30

2.9 Affective model based on the Plutchik's wheel 31

3.1 Temporal line situating the social robots for research applications. 34

3.2 Minerva robot 35

3.3 Kismet robot 36

3.4 PaPeRo robot 36

3.5 iCat robot 37

3.6 Maggie robot 38

3.7 Urbano robot 39

3.8 Quori robot 39

3.9 Mario robot 40

3.10 Temporal line situating the social robots for commercial applications. . . 41

3.11 Paro robot 41

3.12 AIBO robot 42

3.13 Pleo robot 43

3.14 Keepon robot 43

3.15 NAO robot 44
3.16 Jibo robot 45
3.17 Pepper robot 45
3.18 MiRo robot 46
3.19 Sophia robot 47
3.20 Ari robot 48
3.21 Hormonal modulation mechanim proposed by Ávila and Cañamero 49
3.22 ISD architecuture 53

4.1 HRI scenario proposed by Ritschel and colleagues for user adaptation . . 61

5.1 The social robot Mini 74
5.2 The social robot Mini v2 75
5.3 The huggable robot of the social robots group in the UC3M 76
5.4 Gneral view of the robot architecture 78
5.5 The multimodal HRI system 79
5.6 The user adaptation system 83
5.7 Classification of the robot's entertaining activities 86

6.1 The *Motivational model* into our software architecture 89
6.2 Our Motivational model 90
6.3 Neuroendocrine interactions represented in our model 91
6.4 Circadian and ultradian rhythms 97
6.5 ACTH and cortisol episodic ultradian rhythms 98
6.6 Pleasantness-arousal space for determining the robot's mood 104

7.1 The DMS into our software architecture 107
7.2 The decision-making loop proposed in this thesis 109
7.3 Tablet menus to decide the activity 112
7.4 Our DMS architecture 113
7.5 State machine describing the working states of a conditional skill 116
7.6 Communications inside the DMS 118
7.7 Communications of the Central controller with external modules 122

8.1 Front view of the social robot MiRo 136
8.2 The proposed adaptive affective architecture 136
8.3 MiRo homeostatic processes evolution 138
8.4 Evolution of the external signals influencing the IQ 143
8.5 MiRo's affective states 144
8.6 The DMS included in the social robot MiRo 146
8.7 Emotional expressions representing each of the affective states on MiRo. . 148
8.8 MiRo working scenario 150
8.9 Percentage of activation of the robot's motivational states during the non-adaptive trials in simulation 153
8.10 Percentage of activation of the robot's emotional states during the non-adaptive trials in simulation 155
8.11 Evolution of the valence of the robot during the non-adaptive experiment in simulated trials 156
8.12 Percentages of activation of the motivational (left) and emotional (right) states of the robot during the adaptive experiment in simulated domains . 157
8.13 Evolution of the valence of the robot during the adaptive scenario in simulation 157
8.14 Adaptive weights of the internal parameter $\alpha_{internal}$ during the adaptive simulated trials 158
8.15 Percentages of activation of the motivational states during the non-adaptive real experiment 159
8.16 Percentages of activation of the emotional states during the non-adaptive real experiment 160
8.17 Evolution of the valence value during the non-adaptive real experiment . . 161
8.18 Percentages of activation of the motivational and emotional states 162
8.19 Evolution of the valence signal during the real adaptive scenario 163
8.20 Evolution of the $\alpha_{internal}$ weight during the real adaptive experiment 163

9.1 The Preference Learning System 165
9.2 Diagram showing the active user identification and profiling method . . . 167
9.3 SUS question example 170
9.4 Usability ratings for each condition in the user profiling experiment . . . 171
9.5 User engagement at four different moments of the profiling process 172

9.6 Robot entertainment activities organised in hierarchical levels 174
9.7 Data collection and use in the preferences estimation process 175
9.8 Hyperparamenters optimization 179
9.9 Godspeed results in preferences adaptation 182
9.10 Personalisation perceived evaluation for both conditions tested 183
9.11 Set of activities used in the adjustment phase 186
9.12 Diagram showing the dynamics preferences adaptation experiment 189
9.13 Case 1 for optimal preference adaptation 190
9.14 Case 2 for irregular preference adaptations 191
9.15 Adaptive evolution of the Q-value just using indirect user feedback. . . . 191
9.16 Q-value not converging to a solution 192
9.17 Normalised Relative Standard Deviation (RSD) for each condition tested . 193
9.18 Root Mean Squared Error during the preference adaptation 194

10.1 External state of the robot. 200
10.2 Stages of the experiment with Q-learning. 201
10.3 Policy learnt by Q-learning according to the robot's state (S) 203
10.4 Well-being of the robot during the exploration and exploitation phases. . . 204
10.5 Our Dyna-Q+ architecture 209
10.6 Average Mean Squared Error and number of steps to convergence 211
10.7 Evolution of Q-values representing the best action for each robot state . . 212
10.8 Final Q-values learnt by Dyna-Q+ 213
10.9 Evolution of the robot's well-being in the exploration and exploitation phases of the Dyna-Q+ experiment. 214

11.1 Circadian evolution of the *Sleep* and *Wakefulness* biological processes . . 223
11.2 Emotional reactions as a response to dopamine, serotonin, and brain norepinephrine levels 225
11.3 Social behaviour exhibited by the social robot Mini during the interaction 227
11.4 Stress response in the social robot Mini 229
11.5 Long-lasting behaviour exhibited by Mini using the Motivational model when the user is absent 230

11.6 Behaviour exhibited by Mini using the Motivational model when interacting with the user 231
11.7 Evolution of the PB, hormones, stimuli, motivations, and behaviours of the social robot Mini while interacting with a friendly user. 238
11.8 Evolution of the main processes involved in the social response of the social robot Mini during the interaction with an aversive user. 240
11.9 Representation of the bonding and related biological process during the interaction of a neutral user and the social robot Mini. 241

12.1 Moods in the valence-arousal space 245
12.2 Mini's affective state evolution during the long-term interaction 248
12.3 Results regarding the recognition of the mood states of the robot by the participants of the study. 250
12.4 Results of the emotional decay evaluation 251
12.5 Results of the emotional expressions evaluation 251
12.6 Statistical analysis of the four conditions tested in the affective expressiveness experiment 253
12.7 Representation of the signals involved the emulation of the robot's heart rate 258
12.8 Embodied representation of the heart rate in the social robot Mini 258
12.9 Evolution of pupil size during a natural day in the social robot Mini . . . 260
12.10 Representation of the changes in the pupil size of the social robot Mini . . 261
12.11 Blink rate evolution in the social robot Mini 262
12.12 Representation of a blinking sequence in the social robot Mini 262
12.13 Evolution of Mini's breathing rate 263
12.14 Levels of locomotor activity in the social robot Mini 264
12.15 Daily evolution of the locomotor activity in the social robot Mini 265
12.16 Daily evolution of Mini's autonomic functions 266

C.1 Decision-making system UML 321
C.2 Motivational model UML 322

List of Tables

2.1 Summary about the effects of the most important neuroendocrine substances 22

3.1 Summary of the autonomous decision-making architectures based on biological functions reviewed in this contribution. 56

5.1 Comparison of the hardware capabilities of the three social robots developed by our research group. 77

6.1 Interactions between the artificial hormones and neurotransmitters in Mini 92

6.2 Stimuli effects on neuroendocrine substances, somatic, and autonomic processes 95

6.3 Circadian evolution followed by artificial neuroendocrine substances in Mini 96

6.4 Ultradian evolutions of the artificial hormones and neurotrasmitters 97

6.5 Somatic processes included in the Motivational model 99

6.6 Definition of the autonomic physiological processes included in the model 100

6.7 Proactive and reactive motivational states included in the model 102

6.8 Behaviours of the social robot Mini 103

7.1 ManagerControl message example 118

7.2 ManagerFeedback message example 119

7.3 Message from a Manager to control a Skill 120

7.4 Feedback from a Skill to a Manager 121

7.5 Communication between the Motivational model and the Central controller 123

7.6 Example of the message exchanged between the Perception manager and the HRI manager 123

7.7 Example of the message exchanged between the Central controller and the HRI manager 124

7.8 Example of the interaction metrics received by the Central controller . . . 126
7.9 Example of parameters gathered from the Context manager 127
7.10 Example of a message containing a reminder 128
7.11 Example of the message exchanged between the Central controller and the Preference Learning System 129
7.12 Example of the message exchanged between the Central controller and the Behaviour-based Reinforcement Learning System 130

8.1 Homeostatically controlled processes in MiRo 138
8.2 Relationships between the homeostatically controlled processes of the robot, its motivational states, and the stimuli involved in the modulation of the motivational intensities 139
8.3 Definition of the behaviours of the robot during the adaptive social setting 141
8.4 Values of the parameters defining the activation ranges the affective states of the robot 144
8.5 Values of the weights used for modulating the affective behaviour of the robot during the adaptive scenario 147
8.6 Values of the parameters defining the user profiles used in the evaluation of the adaptive system 151
8.7 Values of the parameters defining the robot profiles used during the evaluation of the adaptive system 152

9.1 Results of the PL System validation 180
9.2 Statistical Godspeed results 183
9.3 Statistical results for the Perception perceived attribute 184
9.4 Final results attained by the Q-values preferences learning system 193

10.1 Motivations and biological processes in the Q-learning experiment 199
10.2 Set of behaviours and their effects of the social robot Mini during the Q-learning experiment. 200
10.3 Definition of the biological processes and motivational state in the Dyna-Q+ learning scenario. 206
10.4 List of behaviours of the social robot Mini and their effects during the Dyna-Q+ experiment. 207

11.1 Variables involved in the sleep-wake cycle 220

11.2 Variables involved in Mini's affective response 220

11.3 Variables involved in Mini's social behaviour 220

11.4 Variables involved in Mini's stress response 221

11.5 Effect of PB in how the robot perceives and interprets stimuli 234

11.6 Variables involved in regulating Mini's PB and its social behaviour 235

12.1 Emotions generated by the Motivational model 244

12.2 Kruskal-Wallis tests for the RoSAS attributes warmth, competence, and discomfort regarding comparison of the four conditions tested 252

12.3 Autonomic processes included in Mini 255

D.1 SUS questions 327

Acronyms

ACTH Adrenocorticotropin.

ANE Adrenal norepinephrine.

ANS Autonomic Nervous System.

ASR Automatic Speech Recognition.

AVP Arginine vasopressin.

BNE Brain norepinephrine.

CA Communicative Act.

CART Classification and Regression Trees.

CNS Central Neurvous System.

CRH Corticotropin-releasing hormone.

CT Cortisol.

DA Dopamine.

DL Deep Learning.

DMS Decision-Making System.

EPI Epinephrine.

HPA Hypothalamic-pituitary-adrenal.

HRI Human-robot Interaction.

LED Light Emitting Diode.

LR Label Ranking.

LRF Label Ranking Forests.

MDP Markov Decision Process.

MSE Mean Squared Error.

MT Melatonin.

NLP Natural Language Processing.

OT Oxytocin.

OX Orexin.

PB Pair-bonding.

PL Preference Learning.

PNS Peripheral Nervous System.

RL Reinforcement Learning.

RMSE Root Mean Squared Error.

ROS Robot Operating System.

RPC Ranking by PairWise Comparison.

RSD Relative Standard Deviation.

SAR Social Assistive Robot.

SCN Suprachiasmatic nuclei.

SE Serotonin.

SNS Somatic Nervous System.

SUS System Usability Scale.

TD Temporal Difference.

TLAC Top Label as Class.

TTS Text-to-Speech.

Part I

Introduction

The following chapters introduce the context and motivation of this thesis, emphasising the problem of endowing social robots with biologically inspired decision-making to improve the behaviour they exhibit during long-lasting interaction. Next, we review the existing literature to provide a robust basis about the biology behind human behaviour. Then, we continue this manuscript by a detailed review of the most relevant decision-making systems integrated into social robots, highlighting those based on biologically inspired concepts to generate the robot's autonomous behaviour. Finally, we conclude this part with an overview of current works in modelling adaptive mechanisms and learning models in social robots, focusing on reinforcement learning and preference learning since they are the techniques implemented in our decision-making architecture.

Chapter 1

Introduction

> “My brain is only a receiver, in the Universe there is a core from which we obtain knowledge, strength, and inspiration.”
>
> — Nikola Tesla

1.1. An introduction to decision-making

Curiosity is an inherent feature of human beings. Since childhood, the eagerness for knowledge drives our lives, making us wonder about everything surrounding us. Our desire to learn and improve expands our vision to broader and more complex social contexts. Evolution endowed humans with incredible abilities that any other animal has, especially in interacting with other peers. For this reason, the study of animal behaviour (ethology) has set the spotlight on the neuroscientific basis that shapes human behaviour [1]. Undoubtedly, the human body is an almost perfect organism that depends on many biological processes to maintain the body in good condition. Although relevant advances in the last two centuries have provided a more accurate description of the physiology of the human body, the psychological processes occurring in the human brain have not yet been wholly defined [2]. Therefore, human behaviour is a complex combination of the deficits that our body demands to overcome and the interpretation that each individual makes of the different circumstances they are involved in.

The capacity to produce reasonable decisions and to express how we feel is one of the main human features [3]. Humans make decisions not only giving importance to their physiological needs but performing a complex evaluation of their situation, which has a strong influence on the emotional state of the agent [4].

Researchers attempt to endow machines with autonomous decision-making capabilities considering recent advances in artificial intelligence and robotics. The automation of industrial factories in the past century is the starting point in the rise of autonomous machines. However, nowadays, researchers seek to extend the capabilities of these systems towards more "intelligent" behaviours. Thus, autonomous robotic systems are undergoing an unceasing development for being deployed in many different scenarios [5]. As novel examples in the field it is possible to stand out self-driving vehicles [6], artificial intelligence in video games and simulators [7], strategic military [8], medical applications [9], or robotics [10]. More accurately, if we deepen in the area of social robotics, many challenges have to be overcome to improve the performance of these machines in applications like exhibiting a long-lasting natural behaviour, developing a fluid and appropriate interaction with people, or adapting to dynamic and complex situations without forgetting the principal goal of the robot [11].

Although technology improves by leaps and bounds, a moral debate opens around these applications [12]. Will machines be capable of making autonomous decisions? Should they be conscious about what they are doing? Will robots be able to express emotions? Artificial life and machine learning open a considerable diversity of opportunities for improving our quality of life. However, there are still many problems to solve in the area. This work addresses the main challenges of endowing social robots with autonomous decision-making to exhibit natural and appropriate behaviour while interacting with people and assisting them in their tasks. We focus on designing a system based on biologically inspired concepts that allow the robot to adapt its behaviour to the circumstances of the environment. Decision-making is a broad term that can be defined as "basic cognitive processes of human behaviour by which a preferred option or a course of actions is chosen from among a set of alternatives based on certain criteria" [13]. Thus, this thesis has focused on designing and developing a software autonomous decision-making architecture from a human neuroscientific approach. We believe that for social robots to produce a natural and smooth interaction with people, they should behave similarly to them and adapt to the diverse circumstances of the environment.

1.2. Context

In the last two decades, the social robots laboratory, a division of the Robotics Lab[1] research group at the University Carlos III of Madrid, has focused on researching the application of social robots in Human-robot Interaction (HRI) scenarios. One of the current research lines of the group is designing and developing solutions for older adults with cognitive impairment. Our vision is that robots assist doctors and caregivers in finding an alternative line towards improving the patients' quality of life. At the same time, our robots assist the elders and provide them with companionship and entertainment. Nonetheless, the group also researches in other lines devoted to investigating how to improve the HRI, user adaptation, perception systems, and related applications in social robotics.

Despite the wide range of lines in which our research group works, the origin of the research line behind this thesis arose around 2006, with the development of the

[1] http://roboticslab.uc3m.es/roboticslab/

social robot Maggie [14]. Since then, several PhD theses have seen the light, seeking to research how to improve the interaction with humans and other fundamental aspects of these machines. In 2007, María Malfaz presented her thesis [15] about decision-making based on emotions and self-learning for autonomous social agents. In her work, Malfaz designed a Decision-Making System (DMS) intended for autonomous social agents. The system takes inspiration from biological animal processes to decide how a virtual agent behaves in different situations. The biological basis of the work mathematically represents the agent's needs, motivations, and emotions, important modulators of human behaviour. The agent autonomously manages action selection by considering the state of other agents and resources (objects) in the environment. Using a modification of the classical Reinforcement Learning (RL) algorithm Q-learning [16], the architecture bases the action selection process on a learning environment that considers the artificial deficits of the agent, their effects on its motivational state, and the state of the robot regarding the state of other agents and objects that surround it. The DMS was evaluated in a multi-user virtual environment where the goal of the agent was to maintain the best possible internal state (reduced deficits) while interacting with other agents. During the learning stage, the agent had to explore the environment to obtain enough information about the benefits/problems produced in each situation. Once the agent learned a behaviour policy, the agent exploited it by selecting the best action in each situation. In the virtual world, the agent faces different types of agents (friends, enemies, and neutral), perceiving a wide range of stimuli, whose effects modify the emotional responses of the agent.

Continuing with this research line, Álvaro Castro [17] developed a biologically inspired decision-making architecture for the social robot Maggie [14] ended in 2012. His thesis supposed another step towards endowing social robots with autonomous decision-making since it was implemented and assessed in a real HRI scenario. Like the work developed by Malfaz, the robot's behaviour depends on its motivational and emotional state. Both states depend simultaneously on the deficits that the artificial processes of the robot emulate. Castro considered emotions as discrete independent processes playing an essential role in the motivated behaviour of the robot. In the model, Castro emphasises the role of happiness and sadness on learning, making the robot happier if it receives positive rewards and sadder if the result of an action is unsatisfactory. Using the system, Maggie could perform an autonomous behaviour, including social interactions with people.

Interestingly, the model considers the influence of emotional states on motivated behaviour by studying the role of fear on the fight-flight response to avoid risky situations. The architecture extended the ideas initially presented by Malfaz, endowing a real social robot with autonomous biologically inspired behaviour for interacting with people for a long time. The software was developed to be flexible and modular to be easily implemented in different robots.

Some years later, in a thesis that ended in 2014, Fernando Alonso studied how to improve HRI using multimodal dialogues [18]. Alonso developed a HRI software system for social robots based on Natural Language Processing (NLP) techniques. Initially, the work assessed and tested the most novel algorithms in Automatic Speech Recognition (ASR) and integrated them into the architecture to obtain the best performance of the robot. Once accomplished, Alonso focused on converting text to speech using Text-to-Speech (TTS) motors for allowing the robots to communicate with people verbally. Besides, the system contains non-verbal sounds to improve the expressiveness and naturalness of the robot, allowing it to attain a high level of engagement in the user. The architecture includes robust adaptive mechanisms to personalise the voice and gestures of the robot to different users. The system pretends to enhance the quality of the interaction by considering each person's features that interact with the robot.

Finally, the work started by Pérula [19] in 2017 continued with designing and implementing an autonomous decision-making system in the social robots of the group. The primary contribution of this thesis consisted of including motivational states related to the user's satisfaction and the definition of the robot's skills as functionalities that can be started and stopped at will by the robot, generating a complex sequence of actions. In this context, the robot used Q-learning to map from situations to action to learn how to maintain an optimal welfare state.

The four theses introduced above supposed a solid inspiration for the realisation of this dissertation. In our contribution, we concentrated on extending the previous works, especially autonomous decision-making and biologically inspired models like Malfaz, Castro, and Pérula did. Moreover, drawing on the ideas presented by Alonso, in this thesis, we address user adaptation, including profiling and learning techniques to improve HRI. The architecture developed in this work includes novel advances in neuroscience research, aiming at replicating animal(human) organisms in social robots. The idea is that the autonomous DMS controls the behaviour exhibited by the robot during long-lasting interactions with people. The implementation of the architecture in social robots allows us to investigate the effects of biologically inspired processes in artificial systems, like for example, the role of emotion on behaviour, how deficiencies in biological processes may lead to mental disruption or the effect of biologically inspired behaviour on user acceptance and robot usability.

1.3. Motivation

In the last years, autonomous systems have settled in many social scenarios [20–23]. New trends in artificial intelligence such as Deep Learning (DL) [24] and RL [25] have increased the capabilities of autonomous agents, producing an incredible industrial growth [26].Considering these facts, in the last years, social robots are becoming real in applications such as healthcare [27], education [28], or manufacturing [29].

The deployment of social robots in environments where they cooperate with humans for fulfilling a specific task requires robust HRI mechanisms [30]. Consequently, the social communication between both agents should present a coherent understanding from both sides. In this sense, social robots provide many benefits in their used domains. For example, social robots in healthcare demonstrate valuable outcomes in cognitive stimulation therapies, assistance to caregivers, or social support to elders [31, 32]. Similarly, social robots in education provide positive results as they increase their emotional and cognitive capabilities serving as learning tutors [28]. It is worthy of noting that there exists a great potential of social robots in working with children presenting autism spectrum disorder [33]. Finally, the use of social robots in the service sector has been extended in the last years, facilitating the execution of repetitive tasks and replacing humans where the workforce lacks [34]. Thus, the number of applications where social robots are necessary is increasing. In these contexts, the autonomy of social robots is important to avoid a continuous intervention of human workers in the task [35]. Therefore, for these systems to successfully operate, developers have to endow them with adaptive processes that allow them to modify their behaviour depending on the situation. Thus, to work in increasingly complex scenarios is essential making appropriate, autonomous decisions. The motivation for this thesis arises from the necessity that artificial systems in general, and social robots in particular, have of exhibiting an appropriate autonomous behaviour to attain their goals. Drawing on human capabilities and the way we reason about our actions, decision-making in social robots have taken on a biologically inspired approach for users to feel the robot as a similar peer. In essence, the DMS presented in this manuscript is based on biologically inspired mechanisms to attain naturalness during interaction with people.

Decision-making is, from our point of view, a complex and necessary process that determines the successful operation of the robot. The robot's decisions influence users acceptance and the robot's usability in many different ways, being this influence more notable in social robotics [36]. Recent advances in machine learning and biologically inspired modelling provide researchers with solid foundations to develop more and more novel decision-making architectures for autonomous systems, including social robots [1]. In artificial decision-making, valuable biologically inspired models previously attempted to simulate the natural processes of living beings in autonomous machines in the last years. According to [37], biologically-inspired architectures provide a strong neuroscientific foundation since they ground on how rational animals like humans make decisions. In this work, we pretend to go one step beyond the current works in the literature by deepening in modelling how social behaviour emerges in humans. Hence, we pretend to develop a fully autonomous system that allows our social robots to make appropriate decisions while assisting users with their tasks.

1.4. The problem

The previous section showed how the number of domains where social robots operate has increased in the last decade, introducing the motivation of this thesis. Social robots essentially emerged as an aid to workers for improving people's quality of life [38]. The first social robots presented difficulties adapting to dynamic and unpredictable environments since their goal was predefined and did not allow modifications. Nonetheless, the blossom of artificial intelligence and machine learning in the last years opens a wide range of opportunities and alternatives to endow social robots with the possibility to predict changes in the environment and anticipate them [10]. These improvements permit social robots to improve their capabilities in areas like social adaptation, HRI, learning from past experiences, or exhibiting a natural behaviour. Nevertheless, it is still necessary to continue working in general approaches for social robots to improve their behaviour, especially in fulfilling a natural interaction with their users.

This thesis tackles the problem of developing a decision-making architecture for social robots to overcome three critical issues. First, in complex dynamic environments, robots require a robust DMS that allows them to modify their behaviour to the situation where they are involved, anticipating unexpected changes like the ones that happen in social scenarios. Second, the robot must adapt to the users it interacts with, using past experiences to improve the social communication with the human, driving the task according to their features and preferences. Finally, its users must accept the robot and attain high usability to engage people in long-lasting interactions. Next, we provide our contribution to solving the three problems mentioned before.

1.4.1. Generation of natural behaviours in social robots

A DMS is a module in autonomous architectures that allows the artificial agent to take suitable actions for completing its task. The initial DMSs in artificial systems were based on simple rules that made decisions according to if-else statements using probabilistic rules [39, 40]. The emergence of machine learning techniques and recent advances in neuroscience enabled an important revolution of software architectures that evolved to more complex systems, including adaptive and learning capabilities [41]. Nowadays, many DMSs used in social robotics are based on RL models that represent how the system should act in the environment for completing their task [42–44]. However, these models generally lack in exhibiting a natural, long-lasting adaptive behaviour to the situation and context that allows the agent to react to sudden environmental changes. To overcome the lack of adaptivity and naturalness in the generation of autonomous behaviour, we propose an autonomous DMS based on biologically inspired concepts for making appropriate decisions. The system emulates relevant biological processes in humans, describing physiological and psychological aspects like motivation or emotion that we consider the drivers of the robot's behaviour. Similar works have previously addressed decision-making using biological models for social robots for improving the naturalness with which the agent behaves [45–47]. However, they work in predefined scenarios rather than a complex and dynamic environment like our robots face.

Autonomous decision-makers in real-time domains require robust information to produce appropriate actions. Thus, the perception system is a critical component that provides valuable information for the continuous adaptation of the system. Disturbances in the information handled by the DMS may yield incorrect decisions, leading the system to fail on the fulfilment of its task [48]. In the architecture we present in the following chapters, the decision-making process is supported by a wide diversity of information from the environment, such as the user's features, perceptions from the environment, or deficits in the artificial biological variables that emulate how the robot physiologically and psychologically feels. Using this information, the robot's decisions are expected to become more natural and user-adapted.

1.4.2. Robot adaptation to users

Social relationships are complex processes in the human species. The social features of each person rely on many different factors (e.g. context, preferences, or genetics) that makes it particular of the individual [49]. In this sense, for social robots to accomplish users' expectations and fulfil their tasks, they should exhibit an adaptive social behaviour that personalises the interaction to the user that the robot is facing.

As mentioned in previous sections, social robots can be devised for many different applications. For example, robots used in educative environments require a personalised roadmap for each child [50]. Similarly, social robots dedicated to assisting older adults in cognitive stimulation therapies need customised exercises that have to be conducted explicitly by doctors and therapists. These two examples are just two cases in which user adaptation is essential for engaging users and customising the interaction. In such a way, the DMS might include straightforward mechanisms to identify the users it interacts with and tailor its behaviour and functionalities to them.

So far, the works describing social robots that can modify their behaviour to different users present inflexible mechanisms for changing the interaction mode from one user to another [51, 52]. These mechanisms often require restarting the system to load the user's information or retrain machine learning models when the robot faces a new user. Typically, this lack of flexibility is due to the predefinition of the robot's goal, not allowing it to redefine its goals online [53, 54]. Consequently, to overcome these issues, this thesis proposes an adaptive online method that lets the robot dynamically identify the users it interacts with and appropriately adapt its behaviour to personalise the interaction. Unlike previous works, our method is entirely carried out online, without the necessity to restart the system or modify the robot's hardware settings, allowing the robot to recognise new users and acquire information from them.

1.4.3. User acceptance and robot usability

Although social robots have a great potential for assisting humans in many different tasks [55], recent studies [56, 57] show that most people are reluctant to incorporate robots in their daily lives. The biggest concerns of people about the inclusion of robotic platforms are the replacement of human workers by robots, the belief that a robot is not a capable system (a human can do it better), a sensation of unnaturalness in the interaction and the behaviour the robot exhibits, lack of privacy, and fear to be hurt by the machine. Social robots are significantly affected by the last three cases, provoking many users to feel uncomfortable when using the robot. This fact translates into a lack of confidence and acceptance of the system.

To overcome these problems, we focus on developing a DMS grounded on biologically inspired concepts for making a more capable system that people accept and find easy to use. Users have repeatedly assessed the system to check whether their acceptance and degree usability of the robot were correct. The users' suggestions and the analysis of the features they do not like about the robot allow us to incorporate changes in the robot's behaviour to improve its performance using the mechanisms explained in the previous section.

In this line, from our point of view, the expressiveness and liveliness that the robot shows are very important to transmit naturalness and confidence to the users. For this reason, in this work, we investigate how to represent the affective state and autonomic processes generated by the biologically inspired model using the robot's actuators. Our goal is to transmit to the users how the robot feels for allowing them to understand how the robot is behaving and why it takes a specific action. We believe that exhibiting an appropriate and natural expressiveness may enhance the users' perceptions about the robot, increasing acceptance and engagement in long-lasting interactions.

1.5. Objectives

This thesis aims at developing an autonomous decision-making architecture for social robots to proactively and naturally generate the robot's behaviour.

To reach the main objective of this thesis, additional important sub-goals were defined to achieve the initial requisites of the architecture. We divide the sub-goals into those related to scientific research and those related to the technical aspects.

On the one hand, the goals related to the scientific research carried out in this thesis are the following:

1. As a starting point, we researched modelling techniques of biologically inspired processes for supporting the decisions made by the system. The study comprised the basis of behaviour, motivation, emotion, neuroendocrine responses, or the effect

of the body's deficits on the agent on homeostasis and allostasis, among other terms. Then, we developed a motivational model that embraces these processes, so the robot behaves considering its internal and external situation.

2. An important requirement of the system was to personalise HRI. For this reason, we aimed at developing an active user identification and profiling system to characterise the users that interact with the robot and obtain information from them using interaction. We believe that by personalising the interaction, the robot usability and user acceptance will increase, engaging users. The personalisation of HRI was not carried out in a single stage but divided into several steps that embraced the duration of the thesis. Besides it terms of behaviour adaptation, we seek to endow social robots with a bio-inspired adaptive process that modulates the robot's internal state and behaviour according to the quality and features of the interaction.

3. The robot has to learn from past experiences which actions are more appropriate to its current circumstances. Considering that our robots are designed for entertainment and cognitive stimulation therapies, we focused on learning a general method to map the robot's actions to the situations it is experiencing for maintaining a good internal state. Besides, we concentrated on learning the user's preferences towards the entertainment activities of the robot to personalise HRI. The search for the most optimal algorithm for obtaining such functionality encompassed different stages that led to different implementations and evaluations.

4. Realise a natural and fluid interaction with people translating its biologically inspired processes into expressive reactions based on affect and autonomic control of the robot's actuators.

5. The proposed module had to be validated in real human-robot scenarios to test its operation and versatility. During the assessment, it is fundamental to vary the target population of the different studies that validate the system since the behaviour exhibited by older adults is very different from the one exhibited by children.

On the other hand, the goals related to the operation of the software architecture developed in this thesis are the following.

1. The DMS proposed in this thesis has to be multiplatform so that it can work in all the robotic platforms of the group.

2. The system has to be straightforward to use and understandable by users, even those without any expertise in robotics. The goal is to develop a system that people can use without any problem, allowing the users to fulfil the task they want to execute with the aid of the robot.

3. The system has to be modular to be straightforwardly implemented in the software architecture of the robot. Our robots and other platforms are based not only on

the DMS but also on different modules that must work in synchrony to attain their goals successfully. Thus, for example, it is critical that the DMS correctly communicates with the perception and actuation systems of the robot in order to produce appropriate behaviours.

4. The DMS has to be flexible and adapt to different users. Robot users are broad and diverse, so a machine able to present adaptive capabilities will have more possibilities to be accepted and succeed in accomplishing its goals.

5. The entire system developed for this thesis has to be integrated and run in a real robot. One of the most significant limitations of machine learning algorithms and software architectures is the processing capabilities of the hardware devices that define the physical robot. For this reason, the software must overcome the hardware limitations and be robust enough to carry out its functionality.

1.6. Content overview

The following paragraphs enumerate and describe the organisation of the rest of this thesis.

- **Chapter 2:** Recent advances in neuroscience provide researchers in artificial intelligence with new mechanisms to improve the functionalities of artificial systems. Social robots must present a natural behaviour that humans can easily interpret to meet the people's expectations that emerged during the HRI. This chapter contains the biological foundations of human behaviour from a neuroscientific approach, starting with the functions of simple neurons to the most complex cognitive processes like emotion and motivation.

- **Chapter 3:** This chapter begins by describing necessary research in autonomous social robots. It starts with presenting novel robotic platforms both for research and publicly available. Then, the chapter presents the most relevant and novel decision-making architectures and biologically inspired models in motivated behaviour and affect.

- **Chapter 4:** Social robots act in a complex and dynamic environment. To perceive and adapt to such changes, it is fundamental to endow the robot with machine learning techniques to improve its performance. This chapter introduces research works in learning and adaptation applied to social robotics that inspired our system. Then, we state the learning techniques that we have used during this thesis in terms of Reinforcement Learning (RL) and Preference Learning (PL).

- **Chapter 5:** This work was carried out in the social robots research centre of the RoboticsLab at University Carlos III of Madrid. This chapter presents the principal robots of the group in terms of software and hardware. These platforms implement the software architecture developed in this thesis, extending the robots' functionalities toward endowing them with a logical and natural behaviour. Among the multiple functions (skills) of the robots, we stand out those with a high level of impact on the decision-maker, considering whether they are continuous (working full time) or conditional (discrete activation).

- **Chapter 6:** The biological foundations behind human behaviour were presented in Chapter 2. Drawing on this neuroscientific basis, this chapter presents the motivational model developed for endowing social robots with an internal system that supports the decision made by the robot. The model characterises the evolution of different biological processes (including affective states) as the origin of motivated behaviour and autonomic responses. The output of the motivational model is an essential input of the DMS since it informs the robot about significant deficits in its internal system.

- **Chapter 7:** This chapter describes the autonomous decision-making architecture developed in this thesis. First, the chapter presents the architecture in terms of functionality and operability. Next, the chapter describes how the DMS controls the conditional skills of the robot. Then, we explain how the perception manager of the robot and the agenda manager send information to the DMS modifying its operation depending on the flow of the interaction with people. The DMS considers the motivational and affective state of the robot to make decisions that depend on the intensity of a wide range of internal and external processes. Then, the chapter presents how the biological processes of the robot evolve from a neuroendocrine perspective. Finally, we present essential learning and adaptive mechanisms in the DMS, principally regarding action learning, user profiling, and adaptation preferences.

- **Chapter 8:** In this chapter, we present the evaluation of the robot functionalities related to the experiment related to the stay at the University of Hertfordshire, where we developed an adaptive affective architecture to engage users and personalise the interaction with them.

- **Chapter 9:** In this chapter, we evaluate the Preference Learning System to personalise the selection of entertainment activities to each particular user.

- **Chapter 10:** This chapter describes the Behaviour-based Learning System developed to map situations to behaviours and allow the social robot Mini to learn the best behaviour to execute according to its internal and external state.

- **Chapter 11:** In this chapter, we present experiments related to the biologically inspired behaviour exhibited by the social robot Mini during long-term interactions with people. We analyse the evolution of biological functions emulated in the robot to exhibit natural and autonomous behaviour. Finally, we present an allostatic mechanism to regulate the robot's social behaviour to friendly or aversive users.

- **Chapter 12:** This chapter focuses on introducing the main results regarding the robot expressiveness. It focuses on two approaches: simulating autonomic functions and physically representing them in the robot and assessing the affective expressiveness of our social robot Mini.
- **Chapter 13:** The system developed in this contribution contains some limitations that should be considered. This chapter discusses the main results of this work and presents the limitations providing insights about how to tackle them in the future.
- **Chapter 14:** The dissertation finishes by presenting the main conclusions of this thesis. It also contains future research that will be carried out to improve autonomous decision-making processes in social robots, considering neuroscience, learning, and adaptation research.

CHAPTER 2

Decision-making in humans: biological foundations

The following chapter states the biological foundations that shape the motivational and affective state of the robot as the origin of its behaviour. The decision-making architecture presented in this dissertation attempts to draw on how human behaviour emerges. The following sections continue with a comprehensive overview of the essential processes involved in human behaviour. This review begins with the nervous system as the primary modulator of voluntary and involuntary behaviour. Then, we deepen into the homeostatic and allostatic regulation of biological functions to maintain the body in optimal condition. Next, we move towards a more general vision of how the human body works, presenting its main physiological and psychological processes that can be emulated in social robots. Finally, we describe the importance of motivation and affective states of decision-making and the behaviours we take.

2.1. The nervous system

The nervous system is a collection of specialised cells called neurons that connects the brain and spinal cord with the different parts of our body. Its principal function is to coordinate action and sense the environment to produce appropriate responses [58]. Thus, neurons detect changes in the environment throughout different tissues distributed along the body, reacting according to the intensity of the stimulus. Sensations translate into electric signals that travel to the brain and other body regions, producing coordinated commands to control specific responses [58].

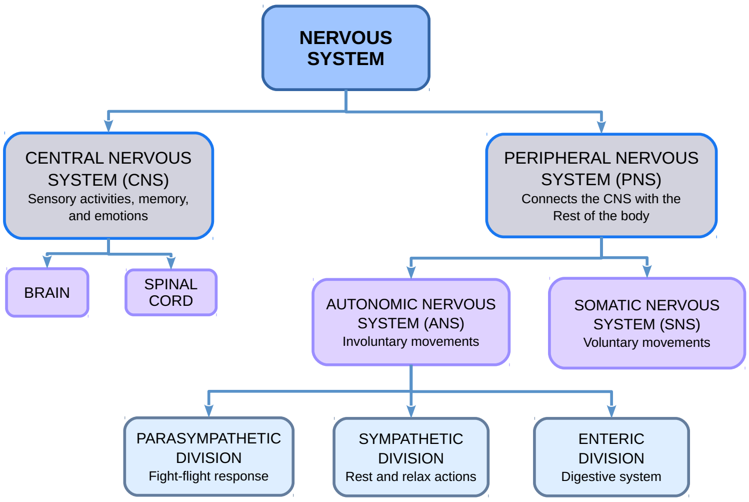


Figure 2.1: Structure of the human nervous system. The nervous system divides into the CNS, formed by the brain and the spinal cord, and the PNS. The PNS subdivides into the ANS, that controls involuntary behaviour, and the SNS, that controls voluntary movements. The ANS divides into three division: parasympathetic, sympathetic, and enteric, controlling autonomic reactions such as heart rate or respiration frequency.

Structurally, the nervous system divides into the Central Neurvous System (CNS) and the Peripheral Nervous System (PNS) (Figure 2.1). On the one hand, the CNS includes the brain and the spinal cord, two important regulators of sensory activities, information processing, and motor control [59]. While the brain is in charge of processing the information of the body sensors and generate appropriate signals for controlling actions, the spinal cord connects the brain to the PNS for transmitting the information [59]. On the other hand, the PNS subdivides into the Autonomic Nervous System (ANS) and the Somatic Nervous System (SNS) [58]. The ANS manages many physiological involuntary processes, like heart rate, temperature regulation, blood pressure, breathing, or pupillary dilation and constriction [60]. Conversely, the SNS controls voluntary movements of the skeletal muscles. The SNS contains afferent nerves that transmit sensations to the CNS, the CNS processes them, and generates motor stimulation carried through efferent nerve fibres in the SNS causing muscle contractions [61].

The ANS divides into the sympathetic division, whose activation leads to high levels of activity and arousal (fight-flight response), the parasympathetic division, in charge of the rest and relax actions, and the enteric division, which controls the digestive system [60]. Like Figure 2.2 shows, increased activity of the sympathetic division causes pupil dilation, increases blinking rate and heart rate, stimulates respiration by dilating the bronchi, and promotes locomotion. On the other hand, the parasympathetic nerves cause the opposite effects [62]. The control exerted by the different divisions of the nervous system over the multiple processes of the body depends on the production and secretion of chemical substances. These substances affect a wide range of biological processes

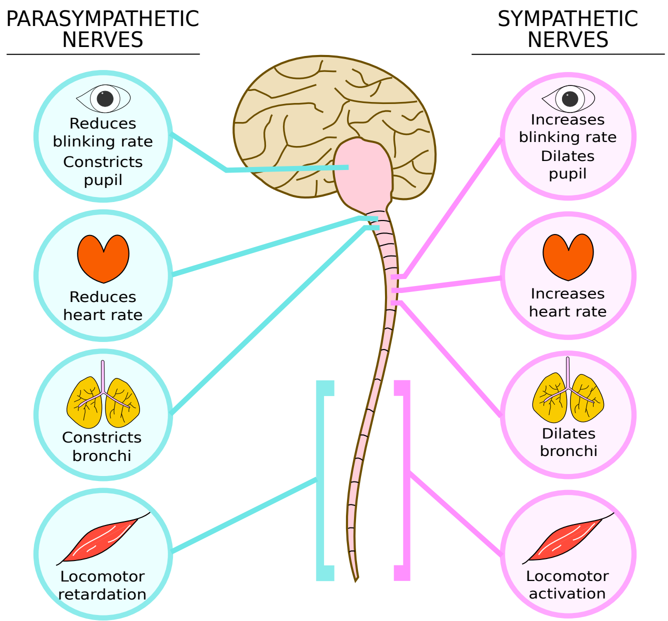


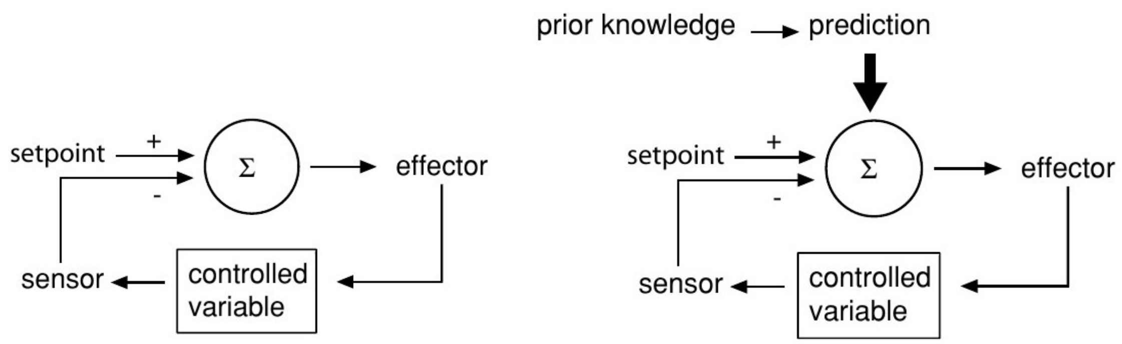
Figure 2.2: Parasympathetic nerves controls the unconscious body actions related to rest. Thus, it reduces eye movement, heart rate, respiration, and locomotion. The activation of the sympathetic nerves leads to the opposite effects: increased blinking, heart, respiration rate, pupil dilation, and locomotion stimulation, among others. These actions arise via the spinal cord, which contains nerve terminals that project to many different regions of the body regulating their rhythms.

[63]. The organism's internal milieu relies on their autonomous regulation and how they regulate our organism. This effect has commonly been denoted as homeostatic control. The following section introduces homeostasis as the maintenance of the internal milieu and allostasis as the adaptation of the homeostatic processes to environmental changes.

2.2. Homeostasis and allostasis

Homeostasis, a term coined in 1927 by Cannon [64] after the novel findings of Bernard [65], refers to the maintenance of the physiological processes of a living organism inside optimal constant ranges by self-regulation. The body's physiological processes must be kept inside optimal ranges to assure the survival of the individual. Deviations from their ideal values cause *deficits* in the physiological state of the agent that must be corrected. Thereby, every physiological process of the organism has an optimal value, and deviations from this value cause a deficit that will lead the agent to correct. Hence, deficits drive our behaviour, motivating us to execute specific actions to maintain our bodies in good condition [66, 67]. An example of a homeostatically controlled variable is body temperature. Under normal conditions, the human body temperature is optimal around 37° Celsius. Besides, it must be above 30° Celsius and below 43° Celsius for survival. If the body temperature is significantly above its optimal value, the body will sweat, eliminating toxins in contact with the air and evaporate and refrigerate the body. Otherwise, if the body temperature is substantially below its optimal value, the body will shiver to produce heat that elevates our body temperature [68]. Extremely abnormal values in a physiological process of an organism can suppose a considerable risk for its survival. Thus, the body has to execute behaviours to correct deficits, not only to maintain good well-being but to survive [65]. It is worthy to note that the ranges and optimal values differ across individuals [69]. It is also important to note that homeostatic regulation affects both voluntary and involuntary processes like thirst or the heartbeat, being the involuntary ones regulated unconsciously.

Although the homeostatic control of the body has been widely accepted in the last decades, it can not explain the variations of some physiological processes regarding environmental changes. For example, the body temperature presents higher values during the late evening and lower values during the sunlight hours [70]. This fact goes against the initial postulate of homeostatic control since not all physiological processes remain stable with time. Considering the previous facts, Schulkin and Sterling defined the term allostasis as the predicted adaptation of the optimal set-point of a physiological process to meet anticipated demand [71]. This definition suggests that regulatory mechanisms are not constant but rather anticipatory to unpredicted conditions. Hence, the homeostatic and allostatic physiological control should be considered as complementary theories instead of opposites, since the allostatic approach extends the homeostatic classical view, like Figure 2.3 depicts.



(a) Homeostatic regulation.

(b) Allostatic regulation.

Figure 2.3: Regulatory mechanisms for the a) homeostatic and b) allostatic control systems. Both approaches are based on a negative feedback loop to correct the error of the controlled variable. Nevertheless, the allostatic control includes a prediction made from the prior knowledge of the system, that is used to minimize the error [72].

Recent studies suggest that the allostatic regulatory model can also be extended to psychological processes, like those involved in social behaviour [73, 74]. For example, our trust or pair-bonding on other people depends on how they behave with us and our own feelings. Thus, the optimal value of these psychological processes varies depending on past and current social experiences. Besides, valuable results invite researchers to believe that psychological circumstances affect the agent's physiology, indicating that physiological processes are also adaptive rather than constant in time [75]. Although an allostatically controlled system is supposed to anticipate errors and, consequently, be more efficient than a homeostatically controlled system, anticipation may be misleading. Thus, allostasis does not guarantee that deficits do not appear but attempts to anticipate them [73]. Allostatic control regulates the organism at multiple levels, from chemical production and secretion to the control of more complex physiological and psychological processes. The following section describes the most critical neuroendocrine responses of the human body, highlighting the influence of neuroendocrine substances in the allostatic control of important biological responses.

2.3. Physiological and psychological aspects of human behaviour

A living organism contains a wide variety of biological processes that simultaneously control the actions of human bodies. These processes actuate together to maintain the internal milieu of the organism. Their disruption may lead to unrepairable deficiencies causing very adverse effects, like, for example, permanent diseases [76]. The following two sections describe the physiological and psychological aspects behind human behaviour, paying particular attention to those that can be implemented in our social robots.

2.3.1. Physiology

Physiology refers to studying the physical and chemical functions occurring in a living organism. It comprises the study of tiny cells to more complex entities like organs [77]. The following sections describe important aspects of human physiology, emphasising the role of neuroendocrine substances in biological functions and how physiological variables are generated and affect human behaviour.

2.3.1.1 The neuroendocrine system

The regulation of the biological functions occurring in the human organism depends on the production and secretion of chemical substances called neuroendocrine substances [78]. Endocrine glands produce and secrete hormones [79] and nerve endings hold neurotransmitters [80]. These substances are influenced by internal and external disturbances of the organism [81]. While neurotransmitters act locally in adjacent cells, hormones influence cells located at different body regions by travelling throughout the bloodstream. Hormones and neurotransmitters have a similar function, altering other cells' operations. It is worth noting that some substances, like dopamine, are considered hormones and neurotransmitters.

Substance	Physiological effects	Psychological effects	Interactions
Melatonin	**Regulates sleep** [82–84] Reduce body temperature [84] **Inhibited by light** [82]	Deficit elicit sad moods [85]	Orexin [86] **Dopamine** [87] CRH [88] **Adrenal norepinephrine** [89]
Orexin	**Regulates wakefulness** [90, 91] Narcolepsy [92] Controls appetite [93] Energy homeostasis [93]	Alertness [94] **Arousal** [94] Attention [93] Memory [93]	Melatonin [86] **Dopamine** [95] **Serotonin** [92] **Brain norepinephrine** [96] **CRH** [97]
Dopamine	Stimulated by light [98] **Locomotor activity** [87, 99] Coordination [99]	Reward seeking [100, 101] **Motivated behaviour** [101] Anhedonia [102], apathy [103], or dysphoria [104] Possible cause of Parkinson [105] **Positive emotion and mood** [106, 107] **Social play motivation** [108]	**Melatonin** [87]
Serotonin	Energy balance [109] **Respiration** [109] **Locomotion** [109]	**Controls negative affective states** [107] Low levels may cause depression [110]	Dopamine [101]
Brain norepinephrine	Fight-flight response [111] **Pupil dilation** [112] **Heart rate** [113] Blood pressure [114] Appetite [115]	**Regulation of anger** and fear [111] Increase attention [116] **Increases arousal** [117] **Social play motivation** [108]	Stress hormones [118] **Dopamine** and serotonin [119]
Oxytocin	**Prosocial behaviour** [120] **Positive physical contact** [121, 122] **Reduces stress** [123, 124] **Pair-bond formation** [125]	Calmed mood [126] Trust [127] Generosity [128] Cooperative behaviour [129]	**Stress hormones** [130]
Arginine vasopressin	**Aggressive behaviour** [131] Blood pressure [132] Circadian rhythms [133, 134]	Controls negative mood disorders (e.g. anxiety, depression) [135]	**Stress hormones** [132] **Dopamine** [133]
Corticotropin-releasing hormone	Food intake [136] **Regulates aggression** [137]	Reduces curiosity [138] Depressive disorders [132]	**Stimulates ACTH** [139]
Adrenocorticotropin	**Similar to CRH**	**Similar to CRH**	**Stimulates cortisol** [130]
Cortisol	**Stress control** [140] Energy homeostasis [141]	Manages anger and fear [140] Territoriality, focused attention, sustained effort, memory, learning, **arousal**, and perception [142]	**Inhibits CRH and ACTH** [130] **Controls norepinephrine and epinephrine levels** [143]
Adrenal epinephrine	**Autonomic functions** **Increases heart rate** [144] Increase blood pressure [145]	**Regulation of anger** and fear (bigger influence on fear) [146] Disruption of this system may cause anxiety and depression [147]	**Stress hormones** [130]
Adrenal norepinephrine	**Similar to epinephrine**	**Similar to epinephrine** but bigger effect on anger	**Similar to epinephrine**

Table 2.1: Summary about the most important physiological and psychological effects caused by each neuroendocrine substance considered in this thesis and the most important interactions among them. In **bold**, those effects simulated in our motivational model.

The following sections describe important neuroendocrine responses occurring in the human body that are included in our motivational model, a module of the DMS that generates the robot's behaviour according to emulated biological functions. Like in humans, Melatonin (MT) and Orexin (OX) control the robot's circadian sleep-wake cycle, that is, the rest-activity phases of the body [148]. The system regulates emotion, motivation, memory, or learning considering the influence from the monoamines Dopamine (DA), Serotonin (SE), and Brain norepinephrine (BNE) [149]. Oxytocin (OT) and Arginine vasopressin (AVP) modulate the robot's social behaviour. While oxytocin controls the positive side of socialisation, arginine vasopressin is involved in regulating aggression [150]. Finally, the Hypothalamic-pituitary-adrenal (HPA) axis clusters the effects of a broad group of hormones such as Corticotropin-releasing hormone (CRH), Arginine vasopressin (AVP), Adrenocorticotropin (ACTH), Cortisol (CT), Epinephrine (EPI), and Adrenal norepinephrine (ANE), that handle the stress response and many autonomic functions of the robot [130]. Table 2.1 summarises the features of these substances on human physiology, psychology, and their most significant interactions with other substances. In **bold**, we emphasise those effects considered in this thesis. Next, we describe the main biological processes (physiological and psychological) affected by neuroendocrine substances that we emulate in our social robots. This description aims to provide a robust biological basis for the motivational model we propose in Chapter 6.

2.3.1.1.1 Sleep-wake cycle and circadian rhythms

Our bodies undergo a wide range of variations during a natural day. These variations, usually called circadian rhythms, control many aspects of our lives such as mood, temperature, or blood pressure [103, 151]. Many studies hold that the circadian pacemaker is in the Suprachiasmatic nuclei (SCN), a brain region in the hypothalamus. Initially, many researchers agreed that circadian rhythms were controlled by the SCN, but recent studies suggest that other glands also participate in the periodic control of important biological processes [130, 144]. Impairments in the SCN cause disruptions in many psychological and physiological processes causing mood disorders like depression [152]. It also causes physical imbalances in body temperature, hormone production, and blood pressure [153]. Recent findings suggest that the SCN reacts to environmental changes anticipating the body responses to, for example, predator risk, food availability, or light conditions [154].

The circadian pacemaker receives abundant stimulation from the retina, which innervates with light [155]. This fact reveals that the sleep-wake cycle is managed by the secretion of melatonin and orexin, two hormones involved in sleep and wakefulness as opposite signals. Melatonin exhibits a circadian rhythm that peaks around 8:00pm with a nadir in the early morning (8:00am) [87]. In daily active animals, melatonin production is interrupted due to unexpected bright light exposure during the night [82], a clear indicator of the strong influence of the light on its endogenous levels. For this reason, melatonin levels inform the body about the timing and duration of the light period [87]. Contrarily, orexin neurons are responsible for wakefulness, alertness, or arousal [93]. The levels of this hormone rise with light, affecting many other processes like shown in Figure 2.4.

2.3.1.1.2 Regulation of affective states

The monoamine nuclei are a group of cells that produce and secrete the monoamine

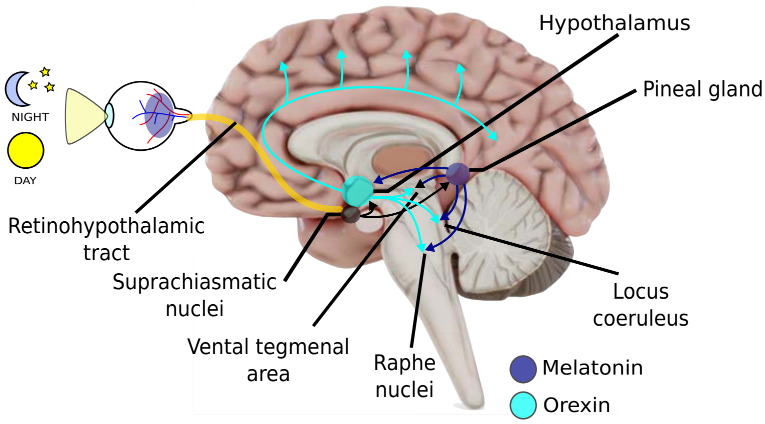


Figure 2.4: Stimulations of the melatonergic and orexinergic regions in the human brain by photic receptors in the retina. The light-dark phase controls the SCN, which regulates the endogenous basal levels of both hormones. While melatonin controls the sleep drive, orexin controls wakefulness.

neurotransmitters dopamine, serotonin, and brain norepinephrine [156]. These substances are produced by the limbic system, a brain region that regulates psychological processes like emotion and mood [157]. As Figure 2.5 shows, dopamine, serotonin, and bran norepinephrine project to multiple brain regions controlling important psychological processes like learning, reward processing, or memory. More precisely, dopamine is implicated in positive affective states like joy or happiness [107, 158]. Similarly, serotonin also controls positive mood, modulating negative affective states [109]. Finally, brain norepinephrine is more implicated in arousal states, controlling affective states like anger or fear [107, 112]. The interrelations of these chemicals support the affective state in humans, and their deficits may cause important affective impairments like apathy or anxiety [105]. Table 2.1 details the effects of these substances, emphasising their role in human affect since they are the primary regulators of the affective state in humans.

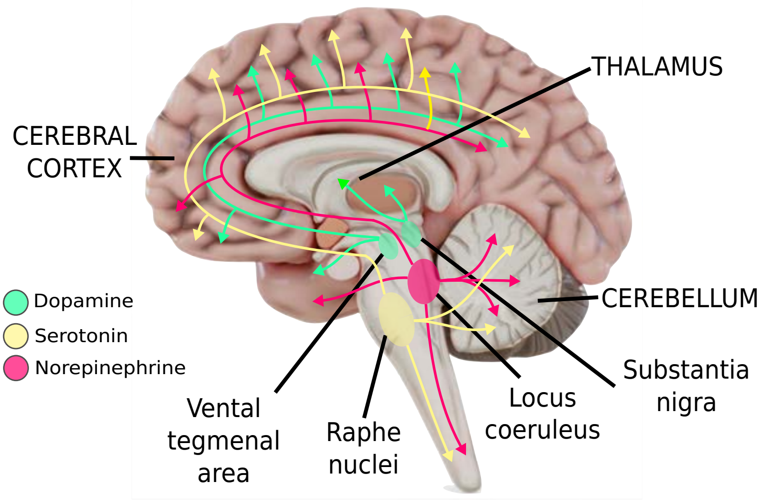


Figure 2.5: The limbic system is the main emotion regulator in the human brain. It clusters the catecholamines dopamine, serotonin, and norepinephrine, whose levels affect many other related processes such as learning, memory, or attention.

2.3.1.1.3 Social relationships

In humans, neuroendocrine substances regulate the most critical processes of the organism, including socialisation [159]. Oxytocin, arginine vasopressin, and CRH are significant modulators of how we relate with other people, being highly affected by the social stimuli that we can perceive from the environment [159]. More precisely, as Figure 2.6 shows, oxytocin is considered the primary regulator of pair-bond formation, prosocial behaviour, and maternal care [120]. Contrarily, arginine vasopressin and CRH are more involved in modulating aggressive behaviour since their levels rise in threatening situations [131]. Table 2.1 details the impact of oxytocin on positive socialisation and the impact of arginine vasopressin and CRH on negative social behaviour like aggression.

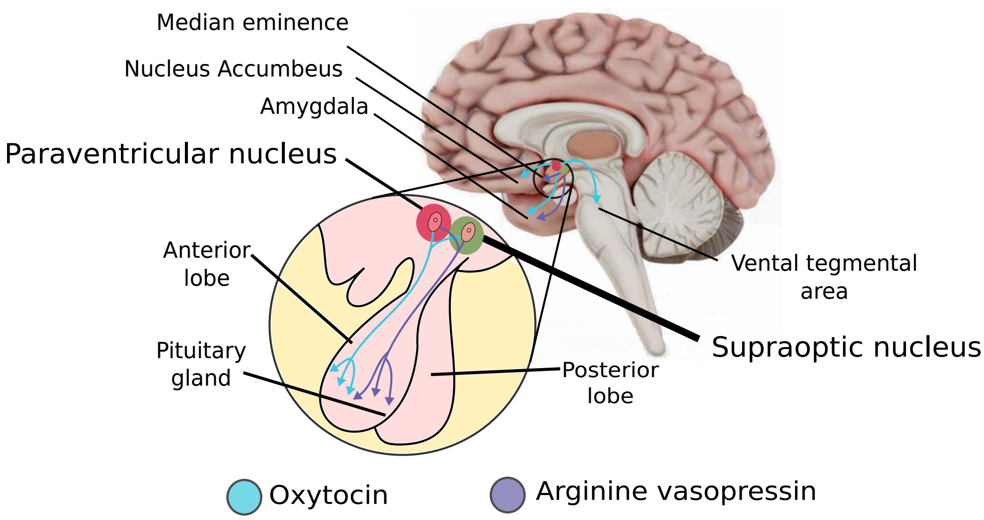


Figure 2.6: The neuropeptides oxytocin and arginine vasopressin actuate at different brain regions principally regulating social behaviour, stress, and emotion.

2.3.1.1.4 The HPA axis

The HPA axis is considered the primary regulator of the stress reactive signal in humans [130]. This response relies on the secretion of different hormones that act synchronously, producing an appropriate reaction to aversive stimuli [160]. Additionally, this system regulates important psychological processes like affective states [161].

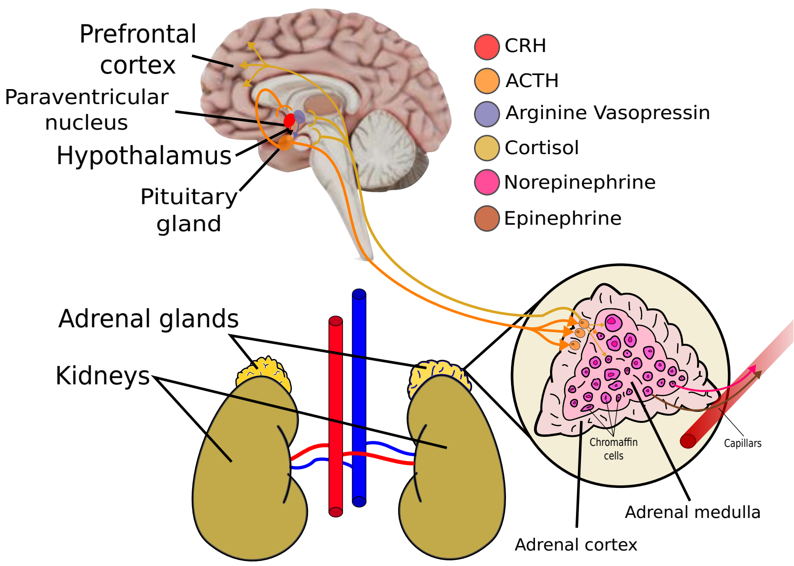


Figure 2.7: The stress response in humans is a complex regulation system that involves many biological processes. The hormones CRH, ACTH, arginine vasopressin, cortisol, epinephrine, and norepinephrine are considered the most important ones, working together to produce an appropriate response to aversive situations.

The stress response in humans begins with the stimulation of the hormones CRH and arginine vasopressin [130]. As Figure 2.7 shows, these hormones are produced in the hypothalamus but stimulate the pituitary gland leading to the production of ACTH [139]. Then, ACTH travels by the bloodstream to the adrenal glands, it drives the production of cortisol, considered the stress hormone [162]. As Table 2.1 shows, cortisol modulates the fight-flight response, promoting the execution of escaping and fighting behaviours in aversive situations [130].

Moreover, significant cortisol amounts translate into epinephrine [143], and norepinephrine, two hormones produced in the adrenal medulla in charge of controlling essential autonomic functions such as heart rate or pupil size [144], as Table 2.1 shows.

2.3.1.2 Physiological functions

Like neuroendocrine substances, high-level physiological functions like sleep or wakefulness follow circadian variations principally controlled by the brain. Some endocrine glands (e.g. adrenal glands) and organs (e.g. heart) also present their internal circadian signal. Like many studies support, the secretion rate of neuroendocrine substances affects physiological processes' evolution. Thus, for example, melatonin is the primary regulator of sleep [83, 163, 164], although other hormones like orexin indirectly influence it [165]. Similarly, orexin is responsible for wakefulness [166], being indirectly controlled by other hormones [148]. Besides, some autonomic functions like heart rate or pupil size also depend on circadian rhythms and norepinephrine/epinephrine levels [167]. Considering these facts, it is possible to assure that the evolution of physiological processes relies on the secretion rate of specific neuroendocrine substances and their circadian rhythms [168].

Considering the wide range of factors affecting physiological processes, the circadian rhythms, neuroendocrine levels, and stimuli are the phenomena we consider for modelling physiological processes in the model we present in Chapters 6. It is worthy of mentioning that these processes vary across individuals since each organism is different from each other [169]. Following an allostatic approach, physiological processes like sleep or heart rate must be maintained inside optimal ranges to guarantee the agent's survival. If these variables deviate from their ideal value, a deficit and a drive will appear, leading the agent to correct it by executing specific behaviours.

2.3.2. Psychology

Unlike physiology, which is distributed in many body regions, psychological processes exclusively occur in the brain [170]. Psychology is highly influenced by the interpretations that each person makes from the stimuli they perceive [171]. These interpretations rely on many different contextual, economic, socio-demographic, and genetic factors [171]. For this reason, the characterisation and modelling of the human psychological process is a complex process since the interpretation of a fact could be across people. Following this approach, cognition depends on a broad diversity of factors like culture, ethnicity or habits [172, 173]. For this reason, modelling psychological processes in artificial agents should be addressed, taking into consideration the previous factors and the tasks that the agent will be devoted to, like Arbib and Fellows discusses in [174].

There are multiple psychological processes in humans, but we focus on describing and modelling emotion, mood, and motivation in this thesis. On the one hand, many theories explain mood and emotion. Important affective theories [175, 176] state that while emotions are reactive responses to unexpected stimuli in the environment by accurate interactions in the limbic system [149], the mood is long-lasting affective states whose activation depend on the accumulation of past and recent experiences [177]. Consequently, both emotions and mood are strongly affected by the stimuli we perceive from the environment and how we interpret them. Whilst emotions are considered reactive responses to certain situations whose effects rapidly disappear, the mood represents the long-term affective state of the agent according to their personality traits and cognitive features [178, 179].

On the other hand, our actions generally depend on our motivational state, representing our physiological deficits and psychological state. Motivated behaviour drives our decision-making and behaviour execution, maintaining our body in good condition.

The following sections deepen in defining the generation of affect in humans considering emotion and mood. Besides, they address how outstanding authors have characterised human affect and motivation in computational models. Next, we present how motivated behaviour emerges depending on the deficits of the organism, the stimuli we perceive and how we interpret the situation we are experiencing.

2.3.2.1 Emotion and mood

As neuroscience advances, the emulation of the psychological process in artificial life follows different approaches. Since the pioneering work of Velásquez [180] about modelling emotion and mood, numerous sophisticated models have been designed. Velásquez research received strong influence from Picard's studies about emotional activation [181] and Tomkins' studies in positive affective generation [182]. The Cathexis model considers cognitive, neural, motivational, and sensorimotor information as the primary elicitors of affect in humans, following Izard's theories [181]. Velásquez distinguishes his model from emotions and mood to differentiate between the individual's short-term and long-lasting affective states. However, his model lacks in explaining the real impact of affective states on decision-making and, consequently, on motivated behaviour. In the last century, important research has been conducted in neuroscience to define how humans' emotional and mood processes arise. Ortony studied in detail the cognitive structure of emotional processing in the human brain [183], concluding that affect depends on a complex evaluation in the human brain dependent on physiological and mental states, but also on subjective and objective evaluations carried out by the individual. Ortony argues that affect generation is a mental process that influences humans' behaviour, cognition, and expressiveness.

Sloman followed a different way in the study of emotions, focusing on developing affective architectures for artificial agents taking inspiration from how humans communicate them to other agents [184]. In his model, Sloman presents a hierarchical system in which different information merges to shape the affective states of the agent. In this case, Sloman includes affective states' effect on motivated behaviour since the artificial system can interrupt ongoing actions if the affective model receives new information. Finally, he concludes the work by discussing whether it is possible to replicate affective communication in artificial systems, which he thought was a difficult task. In line with Sloman, Ekman studied emotions attempting to find a robust relation between human traits and basic emotions. He determined that there exist six basic emotions in human beings: anger, happiness, sadness, surprise, disgust, and fear. According to Ekman, these emotions are independent of sex, or ethnic features [185]. Ekman explored how humans facially express these basic emotions, providing meaningful results about the genetic background in emotion generation and expression [176].

Other approaches have characterised emotion and mood in a more theoretical manner. Thus, Rolls deepen in the definition, nature and functions of emotion in the human brain [186]. In his theory, emotions rely on the reinforcements obtained by the agent while executing actions, the intensity of the reinforcer, and environmental stimuli. According to his beliefs, emotions affect many other behavioural patterns like reward processing, pleasure, or motivation, resulting in complex interrelations of brain processes. Like Rolls, Frijda studied emotion, defining a set of laws that explain how emotions are elicited and influence human behaviour [187]. He posits that emotions are accessible to intentional control to a limited extent, suggesting that we cannot control ourselves in specific situations but are emotions that control us. Frijda was also curious about if endowing artificial systems with emotions was possible [188, 189]. He designed a computer-based model of emotion to react to specific situations (e.g. mistakes, successes, interactions), using animal ethology as the basis of behaviour generation and emotion communication.

Great affective models using dimensions have been designed in the last two decades using the above mentioned and other relevant theories of mood and emotion. Russell's Circumplex Model of Affect [190] places emotions in a bidimensional space, where the valence represents the well-being of the agent and arousal its degree of excitement. The location of each emotion inside the space was carried out using the assessment of users that participate in an experiment in which they indicate how they feel (pleasantness and alertness) when experiencing different emotions. Similarly to Russel, Plutchik [175] used a tridimensional space to characterise emotions. Additionally to the valence and arousal axis, the third axis represents the intensity of the emotion, allowing discrimination among different emotions like anger and rage. More recently, Lövheim developed a three-axis cubic model of emotions based on the levels of monoamines (dopamine, serotonin, and brain norepinephrine). Each emotion is at the vertex of the cube. The level of each monoamine defines a point inside the cube used to calculate the intensity of each emotion. These three models concentrate on the mathematical characterisation of emotion but do not address their impact on other human processes like motivation or decision-making. Additionally, although all these works differentiate mood and emotion in humans, they do not clearly state the beginning and end of each of them.

Following a similar approach, Lövheim [191] designed a three dimensional model of emotion. Each axis of the dimensional space is represented by one of the most significant monoamines (neurotransmitters) involved in emotional processing: dopamine, serotonin, and norepinephrine. The model, shown in Figure 2.8, has a cubic form in which emotions are distributed on the vertexes of the cube, using the eight basic emotions discovered by Tomkins [192]. Thus, the secretion rate of each monoamine determines a point inside the dimensional space that defines the emotional state produced as an outcome. The negative aspect of the model is that despite mentioning that both internal and external stimuli are the primary stimulators of monoamines, it does not explain how stimuli affect the secretion rates of each monoamine. Leukhin et al. [193] validated the model in a computational system to endow robots with intelligent affective decision-making. The NEUCOGAR model (NEUromodulating COGnitive ARchitecture) [194] presented in 2016 uses the L "ovheim's model [191] to define how monoamines affect physiological and psychological processes in an artificial social agent. However, the architecture is only tested in simulation, lacking the expressiveness side that embodied artificial systems require.

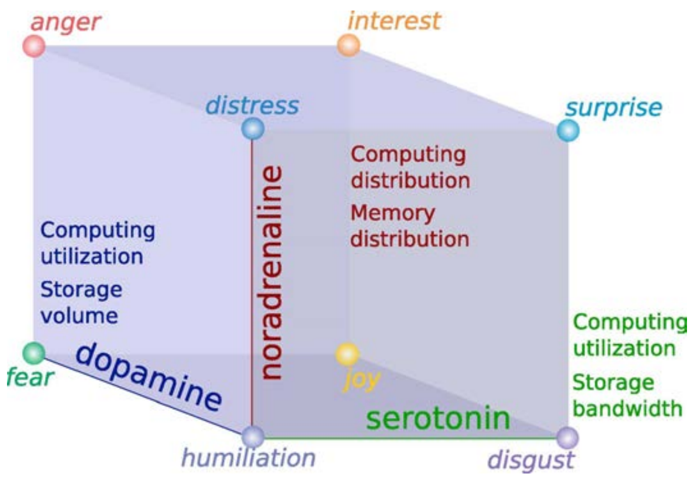


Figure 2.8: Lövheim's three dimensional model of affect [191] based on the secretion of the monoamines dopamine, serotonin, and norepinephrine implemented in the NEUCOGAR model for affect generation [194].

The problem of mood and emotion blending is explored in depth by Zhang et al. [195]. These authors designed in 2016 a tridimensional model of affect that separates emotion, mood, and personality. Besides, they provide mathematical functions to characterise the emotional state, mood state, and agent's personality in each time step. According to them, emotions are people's responses to the confronted events during their interaction with the environment that occur rapidly. Moods represent global emotional states over a long period, being indirectly influenced by the interaction of the environment. Finally, personality is a psychological trait of the individual that defines their emotional tendency. The models integrate emotions, moods, and personalities into a PAD (pleasure-arousal-dominance) cubic space where they are distributed using the results obtained in a user study. Nevertheless, the model does not allow emotional blending since emotions are considered discrete entities without a level of intensity.

The work developed by Qi et al. [196] in 2019 follows Plutchik's approach [175] becoming one of the most outstanding models from our point of view as it includes both mood and emotion in the loop. The model (see Figure 2.9) includes eight basic bipolar emotions and mood: joy-sadness, anger-fear, trust-disgust, and surprise-anticipation. Emotions and moods are continuous entities with an associated level of intensity that vary being affected by different stimuli. For becoming active, the emotional and mood intensities must be above a threshold value that limits the impact of both entities on the affective state of the robot. In the absence of stimuli, the intensity of emotions and mood decays exponentially with time. While the decay rate of moods follows a fixed rate, the emotional decay is related to the intensity of other emotions and for how long it has been active. Interestingly, the model allows emotion and mood overlapping, leading to a more complex and natural affective expression, as we consider in our system.

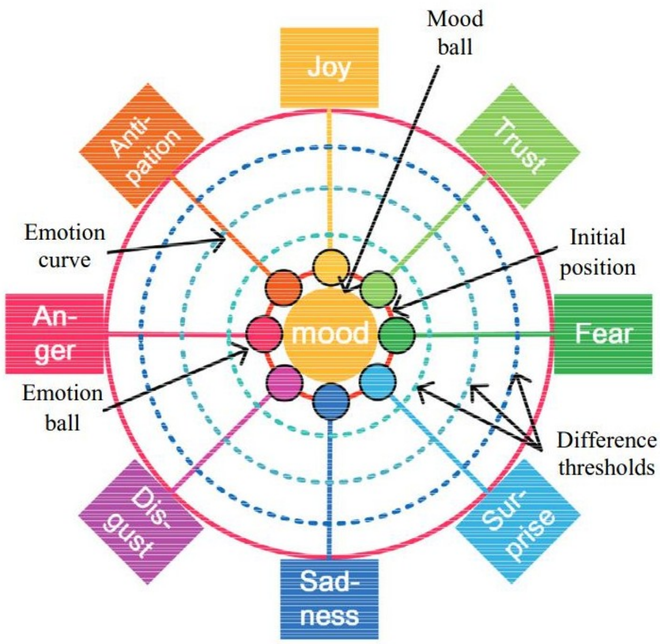


Figure 2.9: Affective model based on the Plutchik's wheel [175] for HRI developed by Qi et al. [196].

2.3.2.2 Motivation

Human behaviour emerges from the physiological and psychological needs of an organism. Motivational states cluster these needs allowing us to evaluate our internal and external state to produce appropriate behaviours. Lorenz described motivations using a hydraulic model where deficits control the flow intensity, and external stimuli amplify their effects [197]. Hence, motivations need deficits to trigger behaviour and stimuli to execute them. For example, humans eat when they are hungry, but they need food to eat and reduce the deficit. Lorenz's model has been used in many artificial systems as the primary gate of behaviour [198–200]. However, this model focuses motivational states on physiology, leaving aside the influence of psychology. Complementary to Lorenz's law, Maslow introduced the concept of hierarchical motivations to discriminate motivational urges during decision-making [201]. This theory organises human needs in a five-levelled pyramid, where physiological needs are in the base, and self-actualisation necessities are on top. Between them, Maslow situates safety, love/belonging, and esteem from bottom to top, suggesting that the agent will reduce physiological needs ahead of safety or social needs in identical conditions.

Using both theories, Velásques [180] posits that motivations arise from physiological needs, but emotions, mood, and related psychological processes also play a crucial role in motivating behaviour. Motivations are therefore affected by a broad diversity of somatic biological processes that define how we decide our actions. McCall and Singer tackle this idea in [202] demonstrating the importance of psychological and cognitive processes on motivation and, therefore, on behaviour. Although voluntary behaviour is usually proactive and involves conscious decision-making, involuntary reactions also take part in motivational states. Like Koolhaas et al. discuss in [203], motivations can be both proactive and reactive depending on whether they elicit conscious or unconscious reactions. Thus, for example, playing a specific game corresponds to proactive motivations while aggressive reactions tally in reactive motivations.

CHAPTER 3

Decision-making in autonomous social robots

This chapter presents outstanding works existing in the literature that inspired this thesis. The first section introduces autonomous social robots for research and those commercially available for the general public. Then, the chapter presents some of the most important biologically inspired decision-making architectures for social robots.

3.1. Autonomous social robots

The following paragraphs present autonomous social robots developed up to today. The review classifies them depending on whether they are used for research or that are available. We focus on describing their most essential functionalities and applications, paying particular attention to their decision-making capabilities and whether they include biologically inspired models to produce such decisions.

3.1.1. Social robots for research

Social robots are dedicated to research areas like human-robot interaction, cognitive science, or healthcare. The review attempts to follow a temporal line, depicted in Figure 3.1, to explain the evolution of these platforms over the years.

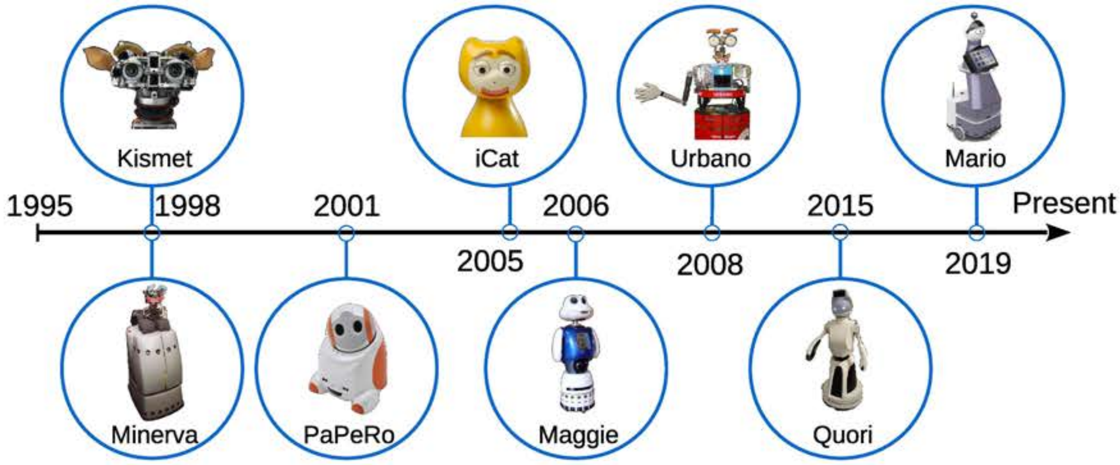


Figure 3.1: Temporal line situating the social robots for research applications.

3.1.1.1 Minerva

In 1998, the mobile social robot Minerva [204] (see Figure 3.2) emerged as the evolution of the tour-guide robot Rhino [205], which is considered as the first social robot working as a tour guide. The robot was deployed in the Smithsonian Museum of American History to guide visitors. Minerva included a navigation system and a Human-robot Interaction (HRI) manager that controls the robot movements in the museum. Besides, the robot has short-term proactive interactions with people. Physically, Minerva improves the capabilities of its predecessor Rhino incorporating an expressive face to communicate its affective state and engage users. It has a differential mobile system combined with an online path generation system based on camera images that allow the robot to precisely localise in the environment. Besides, people could indicate the robot where to navigate in the museum using a web interface that displays the frontal camera of the robot. Therefore, its DMS combined predefined actions indicated by people using the web interface with autonomous actions oriented to guide visitors in the museum. The predefined actions included tours in the museum with a higher priority than the website's teleoperation. Besides, the robot had a reactive layer into its DMS architecture to spontaneously interact with the people it encounters in the museum.

Figure 3.2: Minerva [204].

3.1.1.2 Kismet

Contemporary to Minerva, the social robot Kismet [206] was designed in 1998 to be an artificial creature for interacting physically, affectively, and socially with humans. Kismet, shown in Figure 3.3, aims at eliciting social interactions with people working on many different assistive tasks. The potential of Kismet lies in its anthropomorphic head equipped with 21 degrees of freedom that allow it to orient the head and gaze towards the user. Besides, it can move the eyelids, eyebrows, lips, ears, and many other parts of its expressive face. The robot can perceive users using four colour CCD cameras placed on the front of his face. Kismet understands their vocalisation through a microphone worn by the user during the interaction with people.

The robot's autonomous behaviour depends on a biologically inspired model that defines artificial variables that evolve with time [207]. The architecture has a perception module that retrieves information from the environment and a motivational model representing the robot's needs. Moreover, it has a behaviour module where a DMS decides the following action that is going to execute according to the input information it receives from the perception and motivational model. Then, these abstract actions are translated into separate, specific commands for each actuator. The motivational model combines artificial physiological and emotional processes that are homeostatically controlled. The deficits of these processes motivate the agent to behave in a specific way. Considering the robot's deficits, the behaviour module focuses on reducing the deficits of those processes with a higher level of intensity. Besides, some behaviours require a high deficit to become active and an environmental stimulus to be activated. This autonomous DMS supposed an important step in endowing social robots with natural behaviour taking inspiration from human processes.

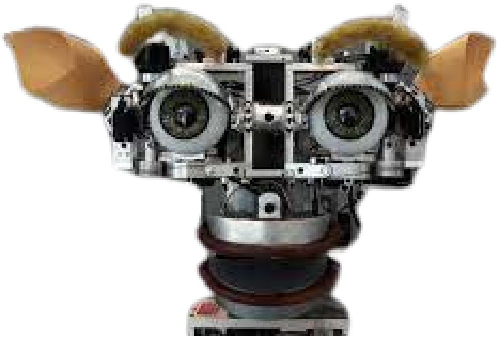


Figure 3.3: Kismet [206].

3.1.1.3 PaPeRo

The Japanese robot PaPeRo [208], depicted in Figure 3.4, appeared as a novel device for childcare. Despite its development started in 1997 with an initial prototype, the first robust version arose in 2001. Initially, it was devised for research, but it became commercially available around 2003. The robot's hardware includes a differential navigation system, ultrasonic receivers located on its chest, ultrasonic obstacle sensors, a hole sensor, and bumpers. The robot can recognise voices using embedded microphones and play sounds using a stereo speaker. By using LEDs placed at its ears, mouth, chest, the robot expresses different emotional states, making it suitable for HRI. PaPeRo can perform a wide range of educational, social, and entertaining applications like telecommunication, playing games, or talking with people. The autonomous behaviour exhibited by PaPeRo combines reactive actions to the stimuli captured by its sensors with planned activities that depend on the audience that the robot is working with and the robot's personality. The personality varied by modulating different aspects of the interaction, like its level of politeness or humanised speech.

Figure 3.4: PaPeRo [208].

3.1.1.4 iCat

The iCat [209] robot, that is shown in Figure 3.5, is a desktop-user interface robot developed in 2005 for HRI research. The hardware system of this robot includes 13 servomotors mainly distributed in its head for performing different affective facial expressions. It has a stereo microphone to recognise the user's speech and the direction of the sound, a build-in webcam to identify human faces, and different touch sensors for perceiving tactile contact. Besides, it incorporates several colours LEDs to improve its expressiveness. The most outstanding functionalities of the robot are its capability to control different home appliances using the home network and the ability to recognise users and create a personalised profile for adapting its domestic functions.

The decision-making system, described in [209], consists of an animation engine that generates the robot's gestures using animation techniques and scripting engines that fusions multiple scripts defining the autonomous behaviour of the robot. Each script represents a robot functionality like talking to the user or planning navigation paths. The system has been designed to allow developers to easily add new functions to the robot, increasing its skills.

Figure 3.5: iCat [209].

3.1.1.5 Maggie

Maggie [14] (see Figure 3.6) was the first social robot developed at our research group, the Robotics Lab from University Carlos III de Madrid, in 2006. It aimed of serving as a robotic platform for HRI research. Over time, Maggie has been used in different applications like gaming [210], translator [211], or home assistant [212]. Maggie can autonomously navigate the environment by using 12 bumpers, 12 infrared sensors, 12 sonar sensors, and a laser scanner. Using an infrared remote controller, it can control home appliances. Maggie can perform verbal communication using ASR and TTS systems, which allow her to react to the user's petitions. The degrees of freedom are in the eyelids, arms, one in the head, one in the neck, and two in a differential wheeled navigation system. It has several LEDs to enhance its expressiveness, touch sensors to perceive tactile contact, and a 3D camera to perform user recognition. Maggie comes with a tablet device used for giving and receiving information to/from the user.

Maggie was one of the first robots to exhibit a fully autonomous behaviour based on biologically inspired artificial processes ruled by deficits and motivations [199]. The architecture, developed by Castro [17] and previously described in Section 1.2, considers motivation and emotions as critical components of the robot's behaviour.

Figure 3.6: Maggie [14].

3.1.1.6 Urbano

The social robot Urbano [213] was developed in 2008 as a museum guide service robot (see Figure 3.7). Its initial DMS allowed the robot to autonomously navigate in the environment and generate appropriate presentations during guided visits to the museum. Some years later, the DMS evolved to a more sophisticated system where decisions are based on the audience's features, the time of the presentation, and other similar factors by grouping predefined sentences containing information about the tour [214]. Using three heuristic search algorithms and fuzzy logic, decisions are made to look for the best quality presentation. First, all the possible paragraph combinations are considered using a 'brute force' technique. Next, a 'greedy best-first' search follows to find a partial optimal solution. Finally, 'backtracking' is performed to find a different alternative if the previous one lacks enough sound quality. Besides, the robot learns from past experiences using the public's feedback to improve its performance in future interactions.

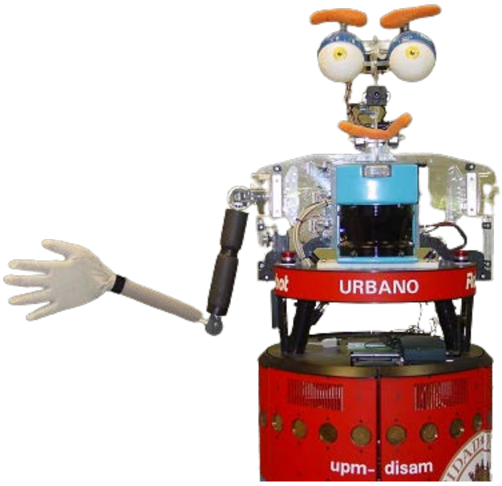


Figure 3.7: Urbano [213].

3.1.1.7 Quori

In 2015, the social robot Quori [215] was conceived for facilitating non-contact human-robot interaction. The upper body is a humanoid-like platform endowed with several degrees of freedom in arms, head, and torso. The robot's base is a holonomic wheel-based system for navigating the environment. Regarding its software architecture, Quori exhibits a fully autonomous navigation and HRI system. The decision-making capabilities of the robot consist of a human-robot dialogue manager that controls the interaction with people and an animal-like perception system for generating coordinated movements of its upper body.

Figure 3.8: Quori [215].

3.1.1.8 Mario

The social robot Mario [216], shown in Figure 3.9, was developed in 2019 as a companion for people living with dementia. Its main application is to reduce the loneliness and isolation of these people by providing them with accessible entertainment applications like the calendar, music, or photos. Mario was born from a European project started in 2016 that aimed at deploying social robots in homes of people with dementia. The robot's design included the impressions of users and the recommendations of experts in mental impairment. The robot's hardware includes a differential navigation system and a tablet device to interact with people. Besides, the robot has an expressive face with several degrees of freedom, a speaker to reproduce its speech, and a microphone to understand the users' commands. Its software architecture [217] enables autonomous behaviours combining an automatic and a deliberative level. On the one hand, the automatic level generates fast actions as a response to the stimuli perceived by the robot. On the other hand, the deliberative level produces the most appropriate behaviour according to the robot's goal. The deliberative process is grounded on a knowledge-based short-term memory that stores structured information about mapping situations to actions depending on the robot's state. The model is based on an object-ontology system that defines how the robot uses its internal resources and the resources from the environment.

Figure 3.9: Mario [216].

3.1.2. Commercial social robots for a general audience

Next, we enumerate some of the most successful robots devoted to the general audience. Some are also used for research, but unlike the social robots presented in the previous section, they are commercially available for other applications. Figure 3.10 situate these robots in a temporal line.

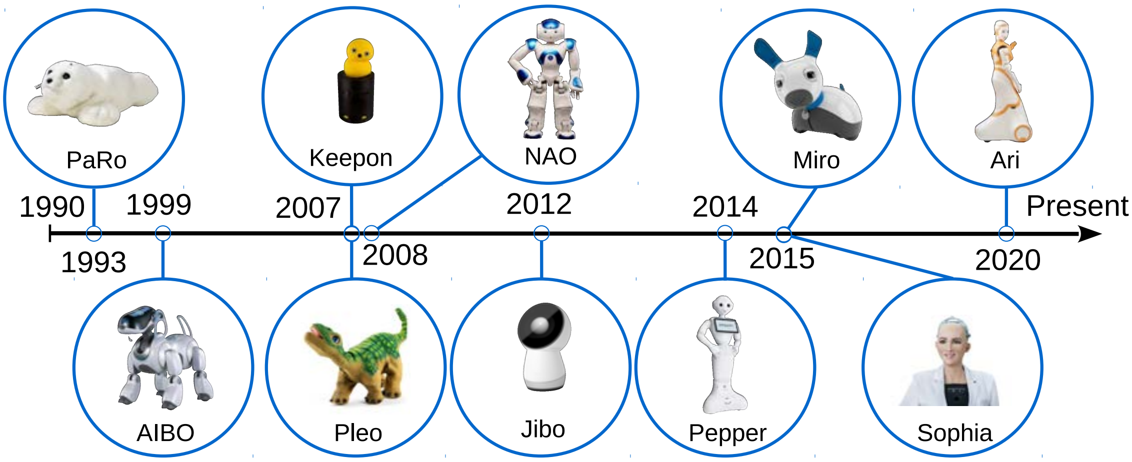


Figure 3.10: Temporal line situating the social robots for commercial applications.

3.1.2.1 PaRo

One of the most successful commercial Social Assistive Robot (SAR) in healthcare is PaRo [218] (see Figure 3.11). Conceived in 1993, PaRo is a shape-like baby seal designed to assist older adults with cognitive impairment and provide companionship. PaRo has an animatronic body that allows it to move the head, the fins, and the tail. It has capacitive sensors in the whiskers and in the body to perceive strokes and a presence sensor in the centre of the forehead to perceive the user's presence. It includes temperature sensors, light sensors to distinguish between light and dark, posture sensors to know when the user holds it, and auditive sensors to improve the perception of the environment. PaRo can identify its partner's voice and react to different utterances using the auditive sensors. The robot exhibits a lively autonomous behaviour that allows it to learn from the interaction with each specific user and adapt to them. Its autonomous behaviour is generated by two hierarchical layers that combine reactive, proactive, and physiological actions. PaRo has demonstrated its benefits in multiple environments assisting older adults with dementia or Alzheimer in elderly residences or hospitals [219–222].

Figure 3.11: Paro [218].

3.1.2.2 AIBO

In 1999, the pet robot AIBO [223] became one of the first commercially available autonomous social robots dedicated to entertainment. The robot, shown in Figure 3.12, presents a sophisticated dog-like shape counting with several sensors and actuators like 16 degrees of freedom, cameras and microphones, or touch sensors. Its first decision-making architecture, called MUTANT, combined deliberative and reflexive behaviour based on biologically inspired processes. The robot's motivations compete among them to rule the robot's behaviour. The robot's motivational states evolve with time depending on the influence of stimuli, simulated instincts inspired from animal behaviour, and emotional reactions that are based on Ekman's model [185].

Some years later, in 2003, Arkin designed a motivational model for AIBO that represents the physiological needs of a dog [198]. The model divides into a drives module that controls the physiological homeostatically controlled variables and an emotional module that generates an intensity value for one of the six basic emotions revealed in Ekman's study [185]. The robot can adopt different attitudes (e.g. dominant, defensive, fighting) that define its behaviour. The robot's attitude is decided by an arbiter and the behaviour by a coordinator that selects the most suitable action from a predefined list. Besides, the motivational intensity depends on specific stimuli that boost motivational intensities, facilitating particular behaviours. Although the model is a simple but effective way of endowing artificial agents with a biologically inspired behaviour, it does not contemplate the interplay between motivation and emotion. This problem was later addressed in [224], a system implemented again in AIBO, where emotions are considered a subset of motivational states that modulate ongoing behavioural control.

Figure 3.12: AIBO [223].

3.1.2.3 Pleo

Pleo [225] is an advanced toy robot used for children entertainment. It became commercially available in 2007. Figure 3.13 shows the appearance of the robot. Pleo is intended to interact with infants during basic leisure activities fostering their creativity, playful learning, and relaxation. The robot is akin to a small rubber dinosaur endowed with 14 motors with customised gears and force-feedback. These capabilities allow it to be a lively pet robot that can perceive its surroundings using a large set of sensors, including capacitive touch perception, a tiny CMOS camera, two microphones, four optical push buttons placed at each foot, two infrared sensors, and a tilt sensor. These capabilities make it suitable for researching in medical and entertainment applications involving sophisticated toy robots interacting with children that may present developmental impairments like autism (e.g. [226–228]).

Figure 3.13: Pleo [225].

3.1.2.4 Keepon

The robot Keepon [229], shown in Figure 3.14, was born in 2007 as a petite robotic chick devised for interacting with children. This robot offers a simple, natural, and non-verbal interaction to aid children with developmental disorders like autism. Its mechanical architecture includes 29 degrees of freedom, although its small size (480mm tall), and several sensors that allow it to be aware of the situation around it, perform facial expressions, and different hand and body gestures. Computationally, Keepon can evaluate situations, mapping them to emotional action while paying attention to the user. The robot can adapt its behaviour depending on the user's actions, presenting suitable reactions to the situation it is experiencing.

Figure 3.14: Keepon [229].

3.1.2.5 NAO

NAO [230], shown in Figure 3.15, is a social robot developed by Aldebaran Robotics in 2008 with great potential in educational and social assistance applications. NAO has been used in various studies with autistic children [231–234] where the robot aimed at promoting engagement during different social and entertainment games. NAO includes many sensors and actuators like microphones for processing the human voice, seven tactile sensors, LEDs, a speaker to play verbal and non-verbal sounds, two 2D cameras, and 25 motors that allow it to walk and manipulate small objects. The commercial version of NAO comes with a software architecture that allows it to exhibit autonomous behaviour. The actions selection consists of a state machine that considers the external environment and the robot's internal state. Thus, its software architecture consists of a perception system that perceives the environment, a control system that generates appropriate behaviours according to the robot's state, an expression system for producing facial expressions and body postures, and a motor system that commands the actuation. Its software, defined as a complex hierarchical state machine, is entirely open so that developers can include their functionalities, expanding the robot's capacities.

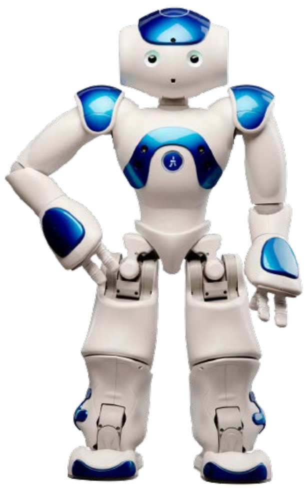


Figure 3.15: NAO [230].

3.1.2.6 Jibo

Jibo [235] appeared in 2012 as one of the first home assistant embodied systems. It became available for purchases in 2017. Figure 3.16 shows a frontal view of the robot Jibo. It is a one-eyed robotic platform designed for HRI in particular homes. Jibo can communicate with users using speech recognition and generation like Alexa or Siri, providing accurate time information about the weather, last news, or available products in superstores. The robot behaviour combines responses to users' questions with proactive communications like reminders.

Figure 3.16: Jibo [235].

3.1.2.7 Pepper

One of the most worldwide used robots in HRI is Pepper [236], a humanoid social robot depicted in Figure 3.17. Peeper was developed by Softbank company in 2014 for being deployed in the service sector. Probably, Pepper is the first autonomous social robot with a sophisticated anthropomorphic design (17 joints), presenting a fully animated head, a very expressive face, two motorised arms for manipulation, and a robust omnidirectional navigation system. Its hardware includes tactile sensors on the hands and head, a tablet device integrated on its chest, 3D and RGB cameras, four microphones, bumpers, sonars, infrared, lasers, and an inertial measurement unit for accurately navigating in the environment. Among its diverse functionalities, it is possible to stand out as a multimodal HRI module that supports a fluid and natural interaction with people. The development kit enables owners to develop new functionalities, basic navigations and manipulation capacities, and lively autonomous behaviour. Up to date, the most important applications of Pepper are entertainment [237], museum guide [238], or education [239].

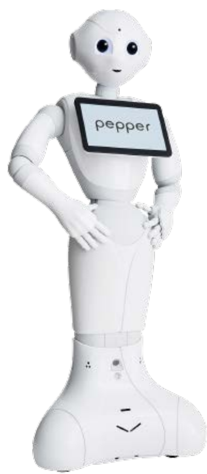


Figure 3.17: Pepper [236].

3.1.2.8 MiRo

MiRo [240] is a social robot that, like its developers define, is intended for engaging the user in science and robotics via edutainment (see Figure 3.18). MiRo emulates a small mammalian pet that can adopt different forms (e.g. dog, rabbit). It was developed in 2015 to work in healthcare (animal therapies), education, and research. The platform has a differential navigation system supported by laser scanners and cliff sensors to wander around the environment. It has touch sensors spread around its shell, light sensors, and luminous devices in different parts of its body. This robot can detect the direction of the sound and respond to it using two microphones located in its ears. Besides, it has several degrees of freedom in the ears, head, neck, eyelids, and tail to perform different emotional expressions.

The behaviour selection of MiRo considers the perceptions from the environment to generate appropriate navigation commands and expressions. Its software architecture is a layered system based on the brain's model that merges reflexive responses to environmental changes with predefined behaviours stored in short-term memory. The system works in a loop fashion, so the decision-makings reevaluate the robot's situation yielding quick motor responses. Thus, the behaviour that MiRo exhibits can be considered fully autonomous biologically inspired behaviour.

Figure 3.18: MiRo [240].

3.1.2.9 Sophia

The humanoid Sophia [241], shown in Figure 3.19, was born in 2015 in Hong Kong, developed by the American company Hanson Robotics. It has the image of a middle-aged woman waist up. The robot can learn how to behave while interacting with people, expressing more than 60 facial expressions, including anger, sadness, or disgust. It can verbally communicate with the user, talking about different topics and communicating meaningful information that can seek on the internet. Sophia can proactively start the conversation thanks to a software architecture based on an autonomous behaviour generation module that integrates the robot's surroundings.

Figure 3.19: Sophia [241].

3.1.2.10 Ari

Finally, the robot Ari [242] (see Figure 3.20) was developed by PAL Robotics in 2020 as an assistive intelligent machine for healthcare applications. The robot integrates novel learning techniques that allow it to learn from repeated interactions with the environment. Moreover, Ari can adapt to the user, presenting a fully personalised behaviour that improves social communication. Physically, Ari has a moving differential base to navigate the environment while avoiding obstacles. A tablet device is used to interact with the user, speech generation and recognition, and two arms manipulate objects. Ari incorporates a software architecture that integrates accurate navigation and localisation, emotional expressions and autonomous behaviour, making it suitable for HRI working as a companion robot. Its autonomous behaviour consists of adapting its navigation and social interaction depending on the stimuli it perceives and its application, such as receptionist registering patients in a hospital, healthcare assistant, entertaining users by executing different activities or conducting rehabilitation therapies by generating the recommended sessions to each patient.

Figure 3.20: Ari [242].

3.2. Biologically inspired decision-making architectures for social robots

Autonomous DMSs allow artificial systems to make appropriate decisions using organised input information. In real-time applications, where decisions typically depend on how the environment changes, the information gathered is essential for accomplishing the tasks. Thus, for social robots to make appropriate decisions, the information they use must be robust and updated.

Since the 1990's and the first years of the 2000's, numerous works have been developed to endow artificial agents with motivated and affective behaviour, [243–247]. This overview starts with the work carried out by Arbib and Fellows in 2004, one of the first works in artificial affect generation for social robots [174, 248, 249]. The authors investigated whether it is possible to include biological processes such as affective states in robots to improve their capabilities and the naturalness in their behaviour. Their works present introductory remarks for endowing social robots with bioinspired processes, focusing on affective states. In this line, Arkin et al. [198] explored in 2003 how to endow animal robots with autonomous behaviour combining deficits, motivations, and affective states. In this work, the biological variables of the robot evolved with time following circadian rhythms, being one of the first works in social robotics to simulate such processes.

In 2005, Ávila and Cañamero [250] initiated an ambitious research line for modelling neuroendocrine responses as the origin of the adapted behaviour exhibited by social robots. The hormonal regulatory mechanism, represented in Figure 3.21, draws on allostasis [72] to modulate the influence of perceptions on motivational states and behaviour. This work goes beyond the previous architectures for generating autonomous behaviour in artificial systems since it explores adaptation from a purely biological perspective. The architecture is tested on real mobile robots, producing good results in the generation of fully autonomous behaviour. As it was one of the first models in including hormonal responses for motivated behaviour, this work supposes a strong inspiration to our DMS.

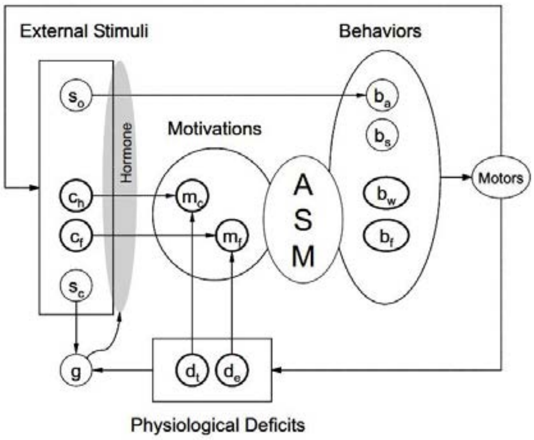


Figure 3.21: Hormonal modulation mechanism proposed by Ávila and Cañamero [250].

Similarly to Ávila and Cañamero, Konidaris and Barto [251] designed in 2006 a motivational architecture for social robots that generates a biologically inspired behaviour. The robot's decisions lie on a drive module representing its deficits and a reward mechanism that lets the agent know how effectively the behaviour is executing. The reward system seeks the correction of its internal deficits. Thus, the model has a module that acknowledges which drive has the highest priority level and a numerical mechanism that links each drive with a specific behaviour depending on past experiences and rewards. Finally, an action selection module decides the agent's action by balancing the drives intensities and the system's priorities. All the artificial variables evolve coordinated to produce a self-maintained autonomous behaviour.

Some years later, Alili et al. [252] designed HATP (for HumanAware Task Planner) in 2009 as a hierarchical task planner for human-robot collaborative scenarios that consider social conventions. The robot's decision-making, called the planning process, relies on two threads in this architecture. On the one hand, the first thread, called refinement, searches for available plans to be executed. On the other hand, the evaluation thread decomposes the decision process into hierarchical sub-problems that can be analysed independently. Finally, the solution that improves the previously considered plan is taken in an iterative process that aims at attaining the most optimal plan. Social rules are context-dependent, producing penalties in the behaviours that do not fit the current context of the social robot.

Contemporary to the previous work, Balkenius et al. developed in 2009 a motivational system for social robots based on the evolution of artificial brain processes [253]. Specifically, the model represents essential functions in the cerebral cortex, amygdala, and orbitofrontal cortex as the origin of human motivation, emotion, and behaviour. The most exciting part of the model is the agent's attention, a mental process that blocks the perception of the stimuli in certain situations leading to the inhibition of the associated reward after a conditioned stimulus. The architecture was tested on a real social robot, allowing an optimal action selection conditioned by the attentional module.

Focusing on the affective component of social behaviour, Samani and Saadatian developed in 2012 an affective architecture that aims at providing robotic systems with affective bonds with the users [254]. The system divides into three modules: the Probabilistic Love Assembly (PLA), a module that calculates the pair-bonding between the robot and the user using a probabilistic model, an Artificial Endocrine System (AES) that includes artificial hormones involved in emotion and physiological needs, and an Affective State Transition (AST) module that controls changes in the robot's affective states using a three-dimensional Cartesian coordinate system defined by motivation, activation, and sub-states. The robot decisions consist of actuation commands that define the robot's expressions, using a neural network that receives information from the three modules mentioned above. From our point of view, the most interesting part of the system is the inclusion of an artificial neuroendocrine system for regulating the affective responses of the robot and the formation of a Pair-bonding (PB) with the user. Both the neuroendocrine system and the human-robot PB suppose an important source of inspiration since these functions are also modelled in our system.

Adopting a different point of view about how to include motivational states in social robots, in 2015, Grigore et al. [255] designed a motivational model for autonomous robot companions. Instead of modelling the robot's motivations, the model estimates the user's motivations and selects appropriate motivational strategies adapting to the user's needs. The robot's strategies are cooperation, competition, self-reflection, and lessons on physical education translated into specific actions that allow the robot to interact with the user. After its execution, each action produces a reward estimated from the user's actions used for learning (using Q-learning) how good the action was. Although motivational strategies suppose a different perspective for action selection, it is not a purely biologically inspired approach, as the model does not consider critical motivational activators like emotions or physiological variables.

Lewis and Cañamero [256] explored in 2016 the role of pleasure on action selection in an autonomous robot. The system contains a motivational module defined by hunger and thirst motivations. The robot can execute behaviour that allows it to look for resources, called appetitive behaviour, and behaviours that allow it to consume resources (eat and drink), called consummatory behaviours. Consummatory behaviours can only be executed if specific resources (food and drink) are available. Otherwise, the robot will have to look for them first. The internal artificial processes of the robot evolve with time, presenting deficits that the robot has to correct. The pleasure mechanisms modulate how likely the robot is to interact with the stimuli in the environment. This factor indicates how much the robot likes each stimulus in the environment, fostering the elicitation of specific behaviour like, for example, preferring palatable food instead of tasteless food.

This work tackles the from of resource availability and how to learn sequences of actions to reduce the robot's internal needs and inspired us for including in our robot's behaviour mechanisms to attain the resources it needs to reduce deficits.

Autonomous social robots must ensure an appropriate social behaviour that considers many vital factors involved in human-human communication. Considering this idea, Cervantes et al. [257] developed in 2016 a decision-making model that integrates the agent's preferences, good and bad previous experiences, ethical rules, and the current emotional state of the agent in the decision process. The model is based on neuroscience, psychology, artificial intelligence, and cognitive informatics research and uses complex mechanisms occurring in the human brain to select appropriate behaviours after evaluating ethical norms and other factors.

Adam et al presents in [258] CAIO, a Cognitive and Affective Interaction-Oriented architecture for social HRI. The architecture was carried out in 2016 for a NAO robot, allowing it to behave while it affectively communicates autonomously. The architecture lies on two fundamental loops: the deliberative and sensorimotor loops. On the one hand, the deliberative loop creates plans and actions that the robot has to execute after evaluating the robot's mental states. On the other hand, the sensorimotor loop integrates the sensors and actuators used for interacting with the environment. Both loops receive information from an emotional module that produces complex emotions from the robot's mental states. The emotional states affect the deliberative loop influencing which actions the robot executes and the motor aspect of the sensorimotor loop defining the expressiveness of the robot. Similarly to the work we present in this dissertation, this architecture combines planned actions with reactive actions, improving the naturalness of the robot. Besides, it is one of the few works that include emotional processes in behaviour selection, enhancing the decision-making capabilities of the robot.

Focusing again on modelling the behaviour of social robots using hormones, Lones et al. [45] studied in 2017 the influence of hormones in the adaptation of homeostatically controlled processes. The system served to define the motivational behaviour of an autonomous social robot that can exhibit a different behavioural pattern depending on the environment where it operates. The architecture includes a motivational module that significantly impacts the robot's decision-making process. Motivational intensities depend on the deficits that the artificial biological processes of the robot and on the perception of stimuli. Each homeostatically controlled variable goes with an artificial hormone whose secretion determines the evolution of the deficits of the robot. The architecture presented in this work is fascinating since it considers hormonal responses and adaptive mechanisms as the origin of motivated behaviour. However, the model does not include the effect of psychological processes on motivations and the robot's expressiveness. Besides, hormonal secretion relies on the evolution of the physiological processes of the robot and a negative feedback loop rather than the stimuli perception. This modelling approach is against previous outstanding studies demonstrating its vast influence on hormonal secretion.

Like the previous works, Schneider et al. [259] continued exploring in 2017 how to shape the motivated behaviour of social robots using learning from demonstration to capture the interaction between a trainee and a trainer. The information learnt by the robot serves as a guideline for future interactions with the user. Then, using previous experiences, the robot creates a finite state machine that contains the motivational actions for accomplishing specific tasks in sportive domains. Motivated behaviour represents simple actions like speaking an utterance to more complex, general actions that involve the participation of the user and the feedback that (s)he provides. Additionally, the model allows synchronising the interaction process with the user, coordinating when each agent takes the initiative.

In 2018, Kowalczuk and Czubenko [260] presented an Intelligent System of Decision-making (ISD) for autonomous social robots based on the mind model. Decisions depend on memory systems, perceptions, representation of knowledge in declarative memory (imagination), attention, learning, and mechanisms involved in decision-making like motivation and emotion. Considering all the previous factors, the agent can select the action that provides it with the best benefits towards maintaining a good physiological and psychological state. This model provides a robust and relevant neuroscientific foundation in the decision process but lacks in establishing the relations between the different processes of the system, like the linking between perceptions and emotion, among others. Besides, it does not deepen the definition of the robot's needs using a homeostatic or allostatic approach but simply uses linear evolution. Although these issues, the architecture, shown in Figure 3.22, supposed a strong inspiration for us because it defines and organises very well the principal biological processes behind human behaviour. The model organises in hierarchical levels where sensory perception (hear, sight, and touch) allows the robot to process the state of the environment, including external objects. Then, the information from the perception is combined with a model of mind where imagination shapes the understanding of the scene the robot is perceiving. Finally, the action selection lies in the robot's needs and affective state (sub-emotions, emotions, and mood). Using a learning module, the robot selects the actions that produce the best benefit for its current states, maintaining its artificial organism in good condition.

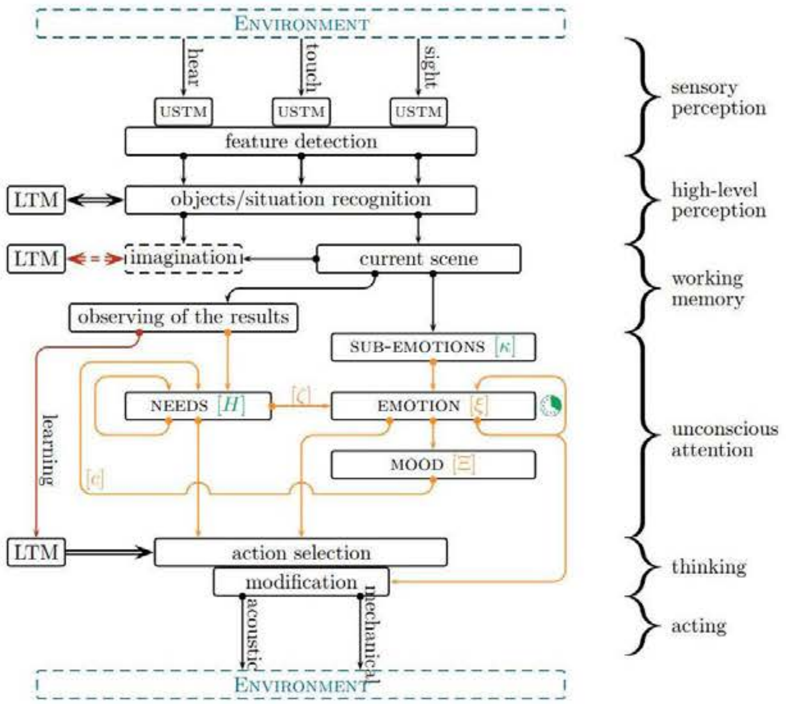


Figure 3.22: General view of the ISD architecture [260].

In 2018, Romero et al. [261] addressed the problem of combining physiological (operational) and psychological (cognitive) drives in the decision-making process of an autonomous social robot. Their system, named MotivEn, contains a utility model that increases the frequency of the robot to achieve its goals by merging perceptions of the environment with drives of the environment. This functionality allows deploying robots in environments with ample state space. At the same time, it provides the robot with adaptive mechanisms to complex and dynamic environments. This work presents a comprehensive distinction between the physiology and psychology of the robot and its influence on behaviour.

In the context of soft social robots, Man and Damasio [46] recently explored in 2019 how to integrate artificial biological homeostatically controlled processes to exhibit natural behaviour. This work intends to provide a framework that allows artificial systems to exhibit autonomous behaviour and be conscious and care about their actions. The artificial processes derive from motivational states that specifically lead the robot to behave. The model also provides the robot with adaptive mechanisms to survive in changing environments, exhibit feelings, and involve consciousness and intelligence in the action selection loop.

Similarly to the previous works, McCall et al. [262] have recently designed in 2019 the cognitive architecture LIDA for cognitive software agents, like social robots. The model combines feelings (including emotions) and motivations to generate a goal-directed behaviour. The architecture combines deliberation and planning in a complex network of modules that considers many factors for deciding how to behave: perception, memory, context, attention, among others. Decisions are translated into motor commands that modify the state of the agent's external environment and provide feedback about the result of the action. Besides, the architecture integrates a RL algorithm for learning which action produces the best benefits in each situation. Interestingly, the system allows general actions like moving or grasping an object and includes more simple behaviours like alarms or events. This work presents many similarities to ours in how it combines RL with a biologically inspired behaviour.

Augello et al. [263] presented in 2020 a purely biologically inspired motivational model for social robots. Drawing upon the somatosensory system of human beings, the model generates the robot's urges according to several factors like its physiological state, competence, certainty, and affiliation, which are influenced by somatosensory information received from the robot's perception system. These urges projects to an action selection module that, using previous experiences stored in short-term memory, leads to the execution of specific behaviours. Like many previous architectures, the system uses RL to define which action produces the best outcomes in the internal state of the robot under each situation. The architecture integrates into a NAO robot whose primary goal is to survive in a dynamic world, maintaining the best possible well-being state, as proposed for our robots.

In a recent attempt to combine motivation and affect, Tanevska et al. [264] developed in 2018 an initial cognitive architecture for the social robot iCub. The novelty of the work is how the robot's affective component changes according to the user's emotional state producing a fully autonomous human-robot expressive interaction. Besides, the robot stores the affective reactions of each user to the behaviour performed by the robot. Then, using this information, the robot can predict the most proper behaviour for each user in the future. During the interaction, the affective module aims at improving both the user's and robot's affective state towards obtaining a good interaction quality. The affective module coordinates with a learning system that allow the robot to anticipate future reactions and learn if the beliefs about each user reaction are true or not. Two years later, the authors [265] updated their previous model supposing a strong inspiration to our DMS. In this second approach, the system's decisions depend on a perception module that gathers information from the environment, an action module that controls the robot's actuation, and an adaptive module that maintains the robot's well-being in good condition. However, the diversity of actions is not too wide for long-term interactions where the user's engagement is a must. Besides, the robot's internal state evolves using linear functions, and even though they simplify the modelling of biological processes, they do not represent the natural evolution that these processes undergo in living organisms.

To conclude with this review, Hong et al. [266] investigated in 2020 the emotional component of affect in autonomous social robots during HRI. The perception system of the robot allows it to perceive the emotional state of the user, adapting its emotional state to be in synchrony with the user. Thus, the authors pretend that their architecture promotes a more natural and engaging communication between both agents. The emotional module consists of a deliberative and a reactive (rule-based) layer. The deliberative layer produces appropriate behaviours and plans, whereas the reactive layer produces instant expressions to respond to certain stimuli in the environment. Behaviours are modulated appropriately to communicate the emotional state of the robot using body language and vocal tuning. Considering its operation, this architecture is similar to ours as it uses deliberation and reactions to drive the decision-making process. Besides, like in our model, emotions modulate the robot's expressions. However, there is a lack in defining the influence of emotions in action selection.

Table 3.1 summarises the previous literature by emphasising the kind of biological processes involved in the robot's decision-making. Most works consider biological functions and the robot's motivational states as the triggers of the robot's autonomous behaviour selection [45, 46, 198, 250–253, 255–257, 259–261, 263, 265–267]. Considering affective states, the most frequent approach is to use emotion as short-term affective states that influence how the robots make decisions and how it expresses [198, 253, 254, 257, 260, 263, 265, 266]. However, very few consider how the robot's mood affects the robot's decision-making and expressiveness as a long-term affective state [257, 260, 266]. Our model considers biological functions as variables that affect motivation, emotion, and mood, defining the robot's internal states. Therefore, the robot's decision-making is affected by the robot's motivations and affective state, defining a robot architecture that includes the most important processes behind human behaviour. Besides, we use emotions and mood to express the robot's affective state to the users, modulating how it behaves. The next section deepens into how the previous works influence our architecture and consider and improve their most important features.

Contribution	Use of physiological or psychological processes	Use of motivations	Use of emotions	Use of mood
Arkin et al. [198]	Emulates the physiology of a dog Circadian rhythm to control arousal	Dominant motivation urges behaviour	Ekman's 6 basic emotions [185]	Not included
Ávila and Cañamero [250]	Hormonal system modulates physiology and motivation	Dominant motivation defines behaviour	Not included	Not included
Konidaris and Barto [251]	Physiological deficits are modelled as drives	Motivations reflect the robot's drives	Not included	Not included
Alili et al. [252]	Not included	Motivations are plans that compete to be executed	Not included	Not included
Balkenius et al. [253]	Models brain regions and their key processes	Motivations are inputs that affect the brain system and action selection	Emotion modulates action selection	Not included
Samani and Saadatian [254]	Complex endocrine model controlling affective states and physiological functions	Not included	Happiness, love, contentment influence decision-making	Not included
Grigore et al. [255]	Not included	Motivational strategies define the robot's behaviour from user actions	Not included	Not included
Lewis and Cañamero [256]	Physiological system hunger and thirst Pleasure hormone to modulate the robot's pleasure to food and drink	Motivation compete to define the robot's behaviour by reflecting its needs	Not included	Not included
Cervantes et al. [257]	Physiological and pleasure systems	Motivations urge behaviour to reduce deficits	Happiness, sadness, anger, fear, and surprise modulate action selection	Considers positive and negative moods
Adam et al. [258]	The robot's state is defined by its physiological and mental state	Not included	Emotions are mental states with influence on decision-making	Not included
Lones et al. [45]	Hormonal responses regulate the robot's energy, health, and temperature	Motivational states urge behaviour from the robot's needs	Not included	Not included
Schneider et al. [259]	Not included	Motivations depends on predefined patters generated from the external state urging the robot's actions	Not included	Not included
Kowalczuk and Czubenko [260]	The mind model Physiological system (needs)	Motivation derives from the robot's emotion, sub-emotions, needs, and mood defining its actions	The model considers emotion and sub-emotion in decision-making	A long term affective factor
Romero et al. [261]	Not included	Motivational states derived from the robot's situation	Not included	Not included
Man and Damasio [46]	Homeostatic model Mental states Feelings	Motivation drives behaviour from the robot's internal state	Not included	Not included
McCall et al. [262]	Basic biological functions	Motivation is derived from the robot's internal needs, external stimuli, and cognitive factors	Emotion and mood	
Augello et al. [263]	Somatosensory system combined with robot's needs, certainty, affiliation, and competence	Deliberative action selection based on motivations derived from the internal system	Vague mention of emotion in the decision loop	Not included
Tanevska et al. [264, 265]	Physiological and psychological systems	Motivational states derived from the robot's needs Action selection oriented to maximise well-being	Happiness, sadness, and neutral modulate the robot's expressiveness	Not included
Hong et al. [266]	A drive/desire system that represents the robot's needs primary during the interaction	Not included	Emotions modulate action selection	The mood affects robot's expressiveness

Table 3.1: Summary of the autonomous decision-making architectures based on biological functions reviewed in this contribution.

3.3. Our contribution

As the previous section shows, the literature regarding autonomous architectures for social robots based on motivation and affect in the last two decades is extensive. Although most of the previous works have influenced our DMS to some extent, our architecture receives strong inspiration from only a few of them.

Among the works mentioned above in autonomous decision-making for social robots, the work developed by Kowalczuk and Czubenko [260] supposed a starting point since it explores and includes in the decision loop many important processes such as memory, perception, motivation, emotion, or learning using a purely biological representation. In our work, we also propose a DMS based on such biological processes to attain a fully autonomous and natural behaviour. In most of the architectures presented above, the DMS concentrates on modelling those aspects involved in the goal they want to fulfil, leaving aside essential factors of the decision-making process like the role of emotion. We believe that for an autonomous agent to be deployed in complex environments and exhibit a fully autonomous intelligent behaviour, the DMS has to integrate both physiological and cognitive inputs. Besides, the work developed by Tanevska et al. [265] in 2020 presents many similarities to our DMS in how it works. Like in our model, the perception system gathers information from the environment and adapts it into understandable information by a decision-maker. Then, the input information is processed by the DMS and translated into the robot's actions. Finally, these actions are adapted depending on different aspects like their emotional state or the features of the user. Our model addresses this idea since the robot's actions are modulated depending on its current affective state and how it evolves with time. However, the model by Tanevska et al. [265] lacks in defining the effect of mood and just uses three emotions.

Considering the modelling of motivational states, in this thesis, motivations include the effect of physiological and psychological deficits and the influence of environmental stimuli [45, 256]. Like in many previous works [46, 198, 207, 256, 261], biological processes are represented as homeostatically controlled variables that evolve with time representing artificial processes in the robot. Drawing on previous research conducted by Cañamero [45, 250, 268], our model considers hormonal responses as the link between environmental stimuli and biological processes (including emotion and mood). The drawback of the previous works is that they are focused on investigating a predefined biological process rather than designing a completely autonomous robot controller for long-lasting interactions. To conclude the motivational aspect, the motivational states of our robot cluster both proactive and reactive motivations as Koolhaas [203] proposes to distinguish between involuntary reactiveness and voluntary decision-making. Thereby, our robot can perform voluntary behaviour while affectively reacting to sudden unexpected events.

In the affective component of behaviour, three works strongly inspired our model. In the first place, the well-known Cathexis framework [180] entailed a complete description of how to endow artificial systems with affective generation and expression. Besides, it explains how to address emotional blending and the different processes that elicit emotional responses using classical but worldwide accepted affective theories. Second, the Lövheim's cube of emotions [191] provides us with a robust background about how to map monoamine levels into emotional states. However, this model lacks in explaining the essential role that mood plays in our affective behaviour. Finally, Arkin et al. [198] modelled the robot's arousal using a circadian signal. However, as we do in this thesis, any work models circadian rhythms to regulate the artificial biological functions and the long-term behaviour of a social robot.

For this reason, we expand our affective vision by including the Zhang et al. [195] three-axis model where emotions, mood, and personality are differentiated from each other, providing a more comprehensive overview of the human affective evolution in time. We propose that the affective model considers emotion and mood for defining the robot's affective state but does not consider personality. The robot's affective state influences decision-making but also modulates the robot's expressions. Finally, the works presented in [194], and [264] gave us a recent and novel way of endowing social robots with affective states. On the one hand, Vallverdú et al. [194] expanded the Lövheims model [191] to a more complex system in which affective states influence the agent's behaviour. On the other hand, Tanevska et al. [264] combined affect and motivation as the origin of the robot's behaviour, presenting outstanding results in real HRI. Thus, as both previous works do, we combine motivation and affect to define the robot's behaviour in our contribution. Additionally, our affective states depend on the artificial secretion rates of neuroendocrine substances, although numerous studies demonstrate that more neuroendocrine substances apart from monoamines also play an essential role in affect [85, 103, 147, 269].

The following chapter reviews the literature for endowing social robots with adaptive and learning mechanisms, focusing on HRI. Then, we present PL and RL, the two learning techniques used in this thesis and how we have included them into our DMS.

CHAPTER 4

Learning and adaptation in social robots

This chapter contains valuable literature about learning and adaptation methods applied to social robots. Then, we describe the techniques we have used for integrating such mechanisms in our architecture, considering Reinforcement Learning (RL) for action learning and Preference Learning (PL) for preference estimation and adaptation.

The software architectures presented in the previous chapter aim at endowing social robots with autonomous behaviour using artificial processes inspired by biology. Although we believe that modelling social robots using human biology allows these systems to exhibit natural behaviour, adaptation and learning in dynamic environments is essential. While adaptation implies that the robot modifies its behaviour depending on the situation that it is experiencing, learning means obtaining new information that is typically used to improve the robot's performance. Next, we present recent literature on learning and adaptation for social robots to operate in variable conditions successfully.

4.1. Learning and adaptation is social robotics

In social robotics, learning and adaptation have been primarily combined for personalising the robot's behaviour to HRI. Following this research line, Ahmad et al. [50] presented in 2017 an adaptive model integrated into a NAO robot for sustaining children entertainment and engagement. In this application, the NAO robot encourages the child to play the well-known snakes and ladders game. The model uses game, emotional, and memory-based online adaptations to avoid the child's boredom. The adaptive game mechanism consists of modifying the verbal communication of the robot depending on the child's actions. The robot's emotional adaptation considers emotional reactions to the inferred emotional state of the children and their actions. Memory-based adaptations consist of storing in a local database and using meaningful information from past experiences in future interactions. Results prove the benefits of the affective and memory-based adaptation of the robot, although no significant outcome is derived from the game adaptive system.

In a similar scenario to the previous work, Schodde et al. [270] proposed in 2017 a method for adapting the communication of a social robot during the interactions with children. The model uses Bayesian probabilistic estimation to drive its decision-making for creating personalised tutoring based on the learner's knowledge. Thus, the robot can generate appropriate sessions to teach each child the skills that (s)he has to learn. The system is implemented on a NAO robot, which conducts the sessions using a tablet device where different multimedia content is played. The robot adapts its communicative procedures during the sessions depending on the dialogue and the children's mistakes using pedagogical acts and expressive non-verbal behaviour.

Using RL, Ritschel and André [271] developed in 2017 a robot architecture based on the adaptation to the user's personality. The rewards obtained by the robot are estimated from the social signals perceived by the user. The adaptation of the robot's behaviour to the human's personality pretends to engage the user. However, in some contexts, the user may prefer a robot with a different personality. For this reason, the RL provides information about whether the user likes or not the actions of the robot, making it possible to change its behaviour appropriately to the desired personality profile. The system works on a real HRI scenario, depicted in Figure 4.1, where the social robot Reeti works as a storyteller varying its degree of extroversion. This work entailed a great inspiration from this thesis since, likewise, we consider explicit and implicit social signals for calculating the users' engagement and how much they like the robot's activities (learning user preferences).

These authors continued working on robot adaptation in applications where the robot acts as a storyteller. In 2018, they [272] investigated how the robot gathers social information from the user. The audio and visual information analysis enables the robot to recognise laughs and smiles as a sign of the user likes what the robot does. Finally, that same year, the authors [273] shaped the humour of a joke teller robot depending on the flow of the interaction. Like in the previous works, social signals are extracted from the user's actions and, using RL, the robot knows whether its actions are appropriate for each situation and whether the user likes them or not. We drawn on for defining how to gather the interaction metrics of the users and learn their preferences towards the robot's actions.

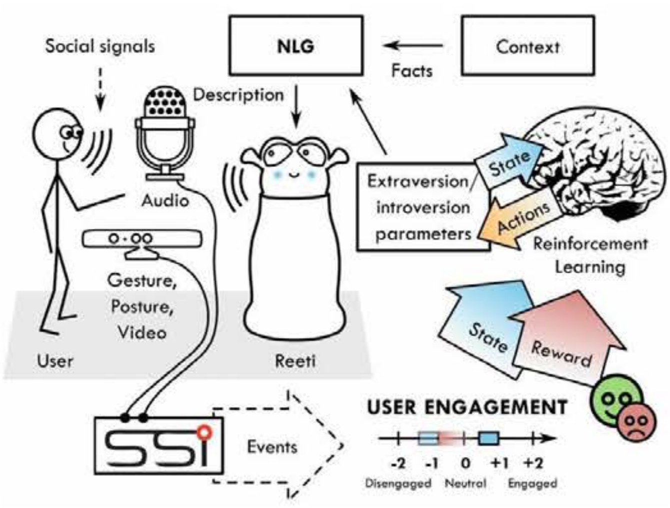


Figure 4.1: HRI scenario proposed by Ritschel and colleagues for user adaptation [271].

Turning the vision to expressiveness adaptation, Chen et al. [274] presented in 2017 a robot that adapts its expressiveness depending on the emotions it detects in the user. For carrying out the behaviour adaptation, the robot uses Q-learning to map emotions to expressions. The robot recognises Ekman's six basic emotions [185], driving its behaviour selection using fuzzy logic. Three robots worked as bartenders in a experiment where 12 volunteers participated. Unlike the previous works, the behaviour adaptation of these robots also including a more complex decision-making by selecting the appropriate drink for each user and their favourite music style.

In the same line, Tanevska et al. [275] researched in 2019 how to improve the affective expressiveness of social robots adapting the robot's personality using a kind of learning similar to RL. The mechanism that was implemented in the iCub robot consisted of the robot extracting facial features and tactile information from the user and adapting its social communication in terms of the duration and frequency of the interaction to maintain engagement. This framework combines adaptive mechanisms inside biologically inspired cognitive architecture modelling processes like motivation and emotion. The adaptive mechanisms are dedicated to the HRI.

The work presented by Andriella et al. [276] in 2020 addressed the problem of deploying robots in wild natural environments for short-term human-robot communication. As the authors argue, a robot that faces inexperienced users may require robust adaptive mechanisms to deal with the unpredictability of the interaction. In the contribution, a TiaGo robot has a cognitive architecture that combines plans with adaptive behaviours using cost functions for defining the actions' suitability. The scenario contemplates the robot interacting with different persons while playing a game. The robot recognises the difficulties of the user adapting its assistance during the game. Results show that the adaptation process leads to a better performance in completing the game, suggesting that an adaptive robot overcomes the user's difficulties.

The work presented in [30] explores how the communication between a robot and a person can be adapted towards obtaining a more natural, fluid, and robust interaction. The work shows the most relevant factors involved in human-robot communication, focusing on exploiting all the robot's physical capacities (sensors and actuators), such as audio, touch, or visual channels.

Finally, Irfan et al. [277] reviewed in 2019 the problem of robot personalisation during long-term interactions. In this work, the authors focus on critical aspects of long-term interactions such as engagement or interaction quality, factors that generally decay with time. The adaptive mechanisms consist of using RL to learn from repeated interactions to map the robot's situation to actions. Thus, the primary goal of the system is personalise HRI to sustain engagement in long-term interactions.

To conclude with this review, we would like to emphasise the work by Rossi et al. [51] as an essential contribution to our adaptive DMS. The review, presented in 2017, addresses how to design adaptive robots for HRI. The authors presented valuable works demonstrating how a social robot can identify the user it is interacting with, create a profile of them, and personalise the interaction using this information. The authors argue that the information should be updated periodically since users' preferences and personal information may change over time. Besides, they posit that the information stored in the robot's memory must contain physical, cognitive, and social data, paying more attention to the ones related to the application of the robot.

Taking a different point of view, Ahmad et al. [278] tackled robot adaptation from a more humanistic perspective. Their contribution focuses on providing a systematic review of how social robots must be designed to operate in healthcare and therapy, education, industrial environments, public spaces, and particular homes. The authors conclude that the future challenges of adaptation in HRI have to be oriented to overcome the adaptation to the users' affective state, personality, culture and the robot's memory and voice.

Considering the previous literature, most contributions combine learning and adapting for improving the behaviour of social robots. Besides, most works are based on RL techniques. The following section defines our approach for improving the adaptive and learning capabilities of our DMS using RL and PL.

4.2. Our approach

The DMS presented in this thesis implements many adaptation and learning mechanisms to improve the behaviour exhibited by the robot. On the one hand, the adaptive mechanisms provide each user with a personalised interaction according to their features. Considering the outstanding review presented by Rossi et al. [51], we opted for defining a unique profile for each user that interacts with the robot. The profile stores different types of information used for customising the interaction, like personal data, habits, or preferences. To allow the robot to create user profiles online and dynamically update the profile of the user that the robot is facing, we designed a novel model based on *Active Learning* that works using face recognition techniques and HRI for retrieving user information. The *Active Learning* profiling system allows the robot to recognise those users that have not interacted with the robot previously, update a face recognition system with their faces, and create them a profile stored in its long-term memory for personalising future interactions with them.

The profile of each user contains their preferences towards the entertainment activities of the robot, so the robot can select the activities that each user likes the most. To develop such functionality, we addressed PL by combining Label Ranking (LR), a promising PL technique with RL for action adaptation, since the goal is to predict the preferences of each user considering their features and refine them using the interaction. Thus, our PL presents the novelty of dividing the process into two stages: first, the LR produces estimations about the preferences of a user based on their features and, with the interaction, adjust the initial predictions using RL. The adaptation of the initial prediction is addressed using social signals provided by the user, like Weber et al. [272, 273] propose.

Finally, and extending the works previously developed in our research group [15, 17] about learning how to map the robot's motivational state to actions, we apply RL techniques to speed up the learning process. Initially, we used Q-learning to deploy such functionality, but due to the slow learning speed of this method, we moved to implement a model-based method like Dyna-Q+ in our social robot.

4.3. Learning methods used in this contribution

In the last years, machine learning applications have obtained meaningful results in many different areas. Among the broad range of machine learning techniques that have been applied to social robotics, we focus on RL and PL. On the one hand, we use RL to allow our robots to map situations to action in order to maintain a good welfare state. On the other hand, we use PL for estimating and adapting the preferences of the user towards the entertainment activities of the robot. Next, we formulate the algorithms used in this contribution.

4.3.1. Reinforcement Learning

RL [25] takes inspiration from how animals learn from rewards and punishments after interacting with the environment. After executing an action, the agent receives a reward that lets them know how beneficial that action was according to its internal and external situation. A RL problem can be formulated as a Markov Decision Process (MDP) that fulfils the Markov's property. This property states that the values of the future variable of a system only depend on present values if Equations 4.1 and 4.2 are analogous. In RL settings, this property is applied to the suitability of executing an action a in a particular state s.

$$Pr\{s_{t+1} = s', r_{t+1} = r | s_t, a_t, r_t, s_{t-1}, a_{t-1}, ..., r_1, s_0, a_0\} \tag{4.1}$$

$$Pr\{s_{t+1} = s', r_{t+1} = r | s_t, a_t\} \tag{4.2}$$

A MDP defines a system modelled using the tuple $\{S, A, T, R, \gamma\}$. In this formulation, S represents a set of discrete state where the agent can be, A represents a set of action that the agent can execute, $T(s, a, s') = P(s`|s, a)$ represents the probability of, after executing an action $a \in A$ in a certain state $s \in S$, transit to a new state $s' \in S$, R a reward function that returns how good is for the agent to execute the action a in the state s, and $\gamma \in (0, 1]$ represents a discount factor applied to minimize the effects of past experiences on new learnt facts. Thus, in a RL scenario, the agent has to learn hot to map situations s_t to actions a_t to accomplish an specific goal. The sequence of behaviours that allow the agent to fulfil its goal is called the optimal policy $\pi : A \times S \rightarrow [0, 1], \pi(a, s) = Pr(a_t = a | s = s_t$.

Numerous algorithms following a RL formulation allow artificial systems to solve learning situations [25]. Nevertheless, it is impossible to describe all of them in this dissertation. Next, we explore the RL algorithms that have been studied in-depth in this thesis, providing a detailed description of their mathematical basis and their advantages and drawbacks. We concentrate on tabular methods since they are suitable for scenarios where the action and state spaces are too large to be saved using a tabular representation.

4.3.1.1 One-step Q-learning

Q-learning [16] is an off-policy tabular algorithm based on temporal difference. It evaluates how beneficial an action turned out to be in a particular state, assigning an optimality value (Q-value) to each state-action pair. Each $Q(s, a)$ is updated when the action a is performed in the state s, eliciting a reward r that increase $Q(s, a)$ if the reward is positive and decreases it otherwise. The update rule of Q-learning depends on α, a learning rate that defines how fast an error δ that considers past and new estimations is corrected (see Equation 4.4). The computational complexity of Q-learning is $O(d)$, as in nature it is a TD standard method [25]. The following equation represents the update rule of Q-learning:

$$Q(s, a) \leftarrow Q(s, a) + \alpha\delta \tag{4.3}$$

where:

$$\delta \leftarrow r + \gamma \max_{a'} Q(s', A) - Q(s, a) \tag{4.4}$$

Despite Q-learning converges to an optimal solution if the reward function is well-defined and if the learning rate α decreases with the number of updates of each Q-value, Q-learning presents some disadvantages. First, it requires a small state-action space since large spaces can not be represented using a table exponentially increases the learning time. Second, in applications where the system can be in danger (self-navigation or self-driving), Q-learning may take unnecessary risks for the system's survival during the learning phase. Finally, Q-learning performs one update per real interaction, so the learning process becomes relatively slow in environments where each action lasts too long.

4.3.1.2 Multi-step Q(λ)

Q-learning provides robust but slow RL algorithms. To overcome this problem, eligibility traces [279] enable that, per time step, the learning system is updated n times instead of once. Thus, Q(λ) emerged as a multi-step updating algorithm that notably speeds up the learning process. Q(λ) considers how recently an action has been executed when updating a Q-value. Then, new actions are executed, the trace value of each (s, a) pair decays, reducing the impact of that action on the Q-value. This mechanism allows the learning system to consider the experience and the n previous experiences in the updating process, updating the past n Q-values. Q(λ) has a computational complexity of $x \cdot O(d)$, where x represents the robot's state space number. Besides, it converges n times faster than Q-learning if both algorithms have the same learning rate and discount factor. Equation 4.5 shows the update rule of Q(λ), and Equation 4.6 how traces are calculated. The traces' decay rate depends on γ, a parameter that determines the impact of past experiences on new updates.

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha\delta_t e(s_t, a_t) \tag{4.5}$$

where:

$$e(s_t, a_t) \leftarrow \begin{cases} e_{t-1}(s, a) + 1 & \text{if a is executed in s} \\ \gamma\lambda e_{t-1}(s, a) & \text{if } Q(s_t, a_t) = \max_{a'} Q(s_{t+1}, a) \\ 0 & \text{otherwise} \end{cases} \tag{4.6}$$

Q(λ) notably improves the learning speed at the cost of increasing the computational charge. However, it is a tabular method, so large state-action spaces are not afforded. Additionally, eligibility traces add extra complexity to the algorithm, making it more difficult to tune it up.

4.3.1.3 Dyna-Q+

Temporal difference methods are model-free, so any environment representation must carry out the agent's learning. However, model-based algorithms have shown notable improvements not only in the learning speed but also in the learning stability [25]. Besides, a model allows simulation while acting, leading to improve action exploration since not regularly visited states in the real domain can be elicited more often using the model.

Dyna-Q+ [280] integrates both real acting and simulation by creating a representative model of the environment. The model uses previous experiences to grow, continuously improving the agent's performance by using the new experiences lived by the agent. In Dyna-Q+, planning starts once the agent has obtained enough real experiences since planning without enough information may lead to mistaken action selection. The potential of Dyna-Q+ resides in the combination of real acting and planning. At the same time, it is based on classical RL algorithms like Q-learning or Q(λ). Humans create their imaginary representation of the environment and create plans about how to behave in the future, reinforcing new behaviours and reducing possible future errors. Dyna-Q+ has a computational complexity of $n \cdot x \cdot O(d)$, where n represents the number of planning steps and x the number of states of the agent. However, Dyna-Q+ requires $n \cdot x$ times more computational processing than Q-learning and n times more computational processing than Q(λ).

Dyna-Q+ speeds up the learning process but increases the computational payload and requires a precise environment representation. Moreover, the complexity of the algorithm is increased regarding classical TD methods. Nevertheless, Dyna-Q+ allows the agent to adapt the learning process to dynamic environments, which Q-learning and Q(λ) can not achieve.

4.3.2. Preference learning

In many moments of our life, we have to make decisions about how to behave. On many occasions, these decisions depend on our physiological needs, but we sometimes select those alternatives according to our preferences. During HRI, the opposite process occurs since the robot has to adapt its behaviour to the user's preferences. PL [281] addresses the problem of allowing artificial systems to estimate the preferences of the user that it is interacting with towards a predefine list of alternatives. Besides, once the estimation is carried out, PL can be combined with RL to tune these predictions, improving the performance of the system. Learning preferences can be tackled using different approaches, depending on the purpose of the application. Following, we describe LR and Preference Adaptation as the algorithms used in this work for learning and adapting the user preferences.

4.3.2.1 Label Ranking

LR [282] predicts, for any instance χ contained in an instance space $\mathcal{X}$, of a preference relation of the type $\succ_\chi \in \mathcal{L}$ among a finite set of labels $\mathcal{L} = \{\lambda_1 ... \lambda_m\}$, where $\lambda_i \succ_\chi \lambda_j$ means that for each instance χ, label λ_i is preferred above label λ_j. Note that the set of labels are ranked according to a total order, defined by a permutation τ_χ of $\{1...m\}$, such $\tau_\chi(i) < \tau_\chi(j)$

whenever $\lambda_i \succ_x \lambda_j$, being $\tau(i)$ the position of λ_i in the ranking and Ω the full permutation space. In this context, we refer to $\mathcal{S}_\Omega$ as the full set of permutations over m labels. Data instances are defined by $(\chi_1, \tau_1)...(\chi_m, \tau_m)$ which contains features and labelled rankings in permutation format. LR aims at mapping from $\mathcal{X} \rightarrow \Omega$, assuming that for every $\tau \in \Omega$ there is a probability $P(\tau|\chi)$ that τ is the permutation associated to instance χ. Next, two of the most widespread PL algorithms that are used in this work are presented [283].

Ranking by PairWise Comparison (RPC) [284] splits the preferences dilemma in binary problems. Considering a finite set of classes $\mathcal{L} = \{\lambda_1...\lambda_m\}$, the system defines $m(m-1)/2$ binary classifiers defined as $\mathcal{M}_{ij}$. For each pair of labels, the classifier establishes a preference relation of the type $\lambda_i > \lambda_j$, indicating that for a specific user, label λ_i is preferred above λ_j. Like Equation 4.7 represents, each model focuses on two labels, but combining all the binary preferences, it is possible to sort all labels in ranking format. The ranking is usually produced using a voting process, ranking labels with more votes above less voted ones. Ties are randomly broken to avoid biasing the classification process.

$$\mathcal{M}_{ij} = \begin{cases} 1 & \text{if} \quad \lambda_i > \lambda_j \\ 0 & \text{otherwise} \end{cases} \tag{4.7}$$

RPC simplifies the classification process at the cost of increasing the computational processing of the algorithm since it trains $m(m-1)/2$ binary classifiers instead of one. Additionally, rank splitting and aggregation are computationally soft, as they rely on straightforward calculus. Despite LR has been widely used in applications like recommender systems [282, 285, 286], just few contributions apply it to social robotics and HRI up to the date [287, 288].

Top Label as Class (TLAC) is a novel LR technique based on replacing the rankings of each input instance by just its top label, which is used as a class [289]. It takes inspiration from Label Ranking Forests (LRF) [290], an algorithm that replaces each ranking with a new label instead of top-ranked labels. TLAC diminishes the number of labels used for classification $\mathcal{L}$ to m classes, instead of $m!$ like [290] does, simplifying the classification problem but requires a preprocessing stage and two rank aggregations that lead to an increase of the computational payload needed for running the preferences estimation. Rank aggregation is performed using the well-known Borda's count, a method that gives a higher score to top-ranked labels to produce the ranking [289, 291, 292]. Although TLAC is computationally faster than similar methods like support vector machines or binary classifiers as it just trains one model for solving the problem, its performance relies on the number of labels to rank [289].

4.3.2.2 Preferences adaptation

Humans feel curiosity about those things that surround them and are known. Sometimes, we have to try things to know if we like them or not. Considering this, in this work we combine RL with PL to adapt the initial user preferences estimations by letting the users try the robot entertainment activities.

Most RL algorithms are based on Markov's property to fully define the agent and his/her steps in the environment. However, there exist methods that do not consider a variable agent's state but learn the most suitable action from a list of possibilities in a fixed state [25]. Action-based methods define an optimality value (Q-value) for action a that belongs to an action space A. The update rule takes inspiration from TD methods, like the following equation represents:

$$\text{NewEstimate} \leftarrow \text{OldEstimate} + \text{StepSize}\left[\underbrace{\text{Target} - \text{OldEstimate}}_{error}\right] \tag{4.8}$$

Every time a new action is taken, their optimality value (NewEstimate) is updated considering its previous optimality value (OldEstimate), a target value (Target) that indicates the new optimality value obtained for the action a and a learning rate (StepSize) that represents how fast the new estimation adapts to the new target. Mathematically, the previous equation can be expressed according to a TD problem by just replacing some of its terms. However, native TD learning uses the agent's states instead of actions to perform the learning, like the following equation represents:

$$V(s) \leftarrow V(s) + \alpha\left[r + \gamma V(s') - V(s)\right] \tag{4.9}$$

In the previous equation, α is the learning rate or step size, r is the reward obtained by the agent, s is the agent's current state, and s' is the next state that will reach after executing an action a.

The adaptation of Equation 4.9 to action-based problems can be quickly made by replacing states values with action values, allowing the comparison of a set of actions instead of states. Thus, fixed learning rates α allow adaptation since the system continuously updates its state.

$$Q(a) \leftarrow Q(a) + \alpha\left[r - Q(a)\right] \tag{4.10}$$

The following Part II describes our system. It start with a presentation of our social robots and their software architecture, ending with the motivational model and DMS we propose to endow our robots with autonomous decision-making.

Part II

The proposed system

This part of the manuscript presents the architecture we have developed for endowing social robots with autonomous decision-making, drawing on how humans make decisions. It is divided into three chapters, oriented to describe the fundamental components of the system:

- **Chapter 5: Our social robots.** This chapter describes the social robots of the Social Robots group at the Robotics Lab, defining their hardware devices and providing a detailed outline of the software architecture where the decision-making system and related modules proposed in this thesis integrate.
- **Chapter 6: Our Motivational model.** In this chapter, we describe the motivational model, one of the most important inputs of the decision-making system entirely developed in this thesis. The motivational model emulates the biological processes of human beings into artificial variables implemented in the robot that define the most appropriate behaviour the robot has to take according to its internal and external state.
- **Chapter 7: Our Decision-making system.** This chapter is the cornerstone of this thesis since it introduces our decision-making system. The chapter details the module's operation to make autonomous decisions and the communications with other modules of the software architecture.

Chapter 5

Our social robots

This chapter begins with a presentation of the social robots where the autonomous DMS has been integrated. Next, the chapter continues presenting the architecture included in these robots. The architecture counts with a set of flexible modules that communicate with the DMS and among them to produce an appropriate and natural robot behaviour.

5.1. Our social robots

Currently, the social robots group works on three different robotic platforms, each one devised for a different application. The social robot Mini, and its new version Mini v2, aims at assisting older adults in cognitive stimulation therapies and provide companionship and entertainment. Both robots have similar capacities, but the reduced size and cost of Mini v2 make it more affordable and easy to use in particular homes. Thus, Mini is typically devoted to research and Mini v2 is intended for performing experiments outside the university. Moreover, the research group is working on the design and development of a huggable social robot. Although it is under development, we pretend that, in the future, we can implement a version of our DMS presented in this thesis in scenarios that use social robots for affective stimulation and companionship emulating animal therapy. The following two sections describe in more detail our social robots for which the DMS has been developed, focusing on its hardware capabilities.

5.1.1. Mini

The social robot Mini [293] (see Figure 5.1) is a robotic platform developed in 2014 intended for cognitive and affective stimulation with older adults to prevent the effects of cognitive impairment. Furthermore, Mini provides older adults with entertainment since it contains a repertoire of different activities like games, multimedia content, informing about the latest news, or the weather forecast.

Figure 5.1: Front view of the social robot Mini and the tablet device attached to it.

Physically, the robot has five degrees of freedom to move its hip, both arms, the neck, and the head using five servomotors. These servomotors communicate with a controller specifically integrated to command them. To improve its expressiveness, Mini has RGB LEDs distributed throughout its body. One of these LEDs emulates the heartbeat of the robot, allowing it to express different affective states. The LEDs located at its cheeks simulate the flushing that, for example, occur during embarrassing situations. Besides, the robot has two animated screens that simulate the robot eyes and two temperature sensors to measure its internal temperature and the environment's temperature. For controlling every actuation device and holding the software architecture that is described in Section 5.2, the robot has a Mini ITX motherboard containing an Intel i7 processor. Additionally, an Arduino MEGA 2560 connected to the motherboard manages the RGB LEDs and the temperature sensors.

Mini can perceive tactile contact using four capacitive touch sensors placed on both shoulders, belly, and head. It is worthy to note that the most recent versions of Mini replace the capacitive touch sensors with contact microphones. Contact microphones allow the robot to differentiate strokes, taps, slaps, and tickles using machine learning techniques, improving the perception system of the robot [294]. In this thesis, we have used a light sensor for measuring the ambient light intensity so that the robot can differentiate between day-night conditions. These sensors are all connected to the Arduino MEGA 2560.

Mini can verbally communicate with people using a TTS system that translates text to speech and a ASR module that translates the user's speech into text for the robot to understand the user's commands. The user's speech is obtained using a microphone with active noise cancellation. Then, a stereo speaker plays the audio created by the TTS system verbally generating the robot's speech. Moreover, for those cases in which the user may present verbal limitations, the robot has a tablet device to show visual and auditive information. This device permits displaying menus on it, so users can select the option they prefer from a predefined list of alternatives. Besides, the robot has a 3D stereo camera attached to the joint placed on its hip in applications that require image processing.

Recently, in 2020, the social robot Mini was updated to a new version [295] aiming at improving its computational capabilities and reducing its base. Mini v2 has the same degrees of freedom as Mini, but their servomotors are cheaper. Instead of including an Arduino MEGA 2560, Mini v2 has an ESP32 to control these devices. The Arduino and the motor controller connect to an Intel NUC motherboard containing an Intel i7 processor. Unlike Mini, this robot integrates a passive infrared sensor that detects presence. It has an array of microphones placed on its head for perceiving the direction of the ambient noise. Besides, Mini v2 incorporates a Raspberry Pi 3 for controlling two sophisticated animated screens that simulate the robot's eyes. These screens extend the functionalities of the Mini's eyes, including the control of pupil size and iris colour. The robot has the same microphone, speaker, and tablet device to verbally communicate with people.

Figure 5.2 shows the frontal view of the social robot Mini v2. As can be perceived from the comparison of the frontal view of Mini v2 and Mini (Figure 5.1), the base that bears the robot is substantially smaller, mainly due to the change of some hardware devices from the base to the torso.

Figure 5.2: The social robot Mini v2.

5.1.2. Huggable robot

The social robots group is currently immersed in developing a new robotic platform devised to provide older adults with companionship trying to emulate animal therapy. Considering outstanding results in medicine using animals to improve the health of patients [296–298], roboticists create pet robots that serve as companions for vulnerable sectors of society like older adults [299]. Using robots instead of animals can help overcome well-known problems like excessive noise, cleaning, or allergies that real pets have. These drawbacks are even more notable when patients are in hospitals or shared-houses.

After studying the possible external appearances of the huggable robot, we opted for selecting the alternative shown in Figure 5.3. The robot is a small and soft rabbit. Thus, the robot can be easily manipulated by older adults, as it has a foamy case so people can hug and stroke it. The robot has capacitive touch sensors spread around its body to perceive the tactile contact, temperature sensors, and a microphone to perceive the human voice. Considering its motor capacities, it has five servomotors for moving both ears, the tail, the neck, and the head.

Moreover, the robot counts with a speaker for playing non-verbal sounds and RGB LEDs that simulate different affective states. Excluding the servomotors handled by their manager device, a SP32 microcontroller handles the other sensory and actuation devices. The servomotors microcontroller and the Arduino are connected to a Raspberry Pi 3 that is in charge of processing the high levels computations of the robot.

Since the robot is portable and not permanently connected to the electric current, it incorporates a 3s-1p LiPo battery that provides an energetic autonomy of approximately 8 hours. Besides, an uninterruptible power supply device manages the charging of the battery when the robot is connected to the alternate current avoiding undesired shuts-off. Although this robot has a different application from the previous ones, it includes a simplified version of the robotic architecture and DMS described in Section 5.2.

Figure 5.3: The huggable robot of the social robots group in the UC3M.

The following Table 5.1 compares the hardware devices integrated into each of our social robots.

5.2. Software architecture

The software architecture developed in our research group uses the well-known Robot Operating System (ROS)[2] software. ROS has multiple advantages like flexibility, scalability, and modularity. ROS allows isolating each robot functionality in a different package/project without affecting the whole architecture, providing robust and straightforward mechanisms to communicate each function. Conceptually, ROS

[2] https://www.ros.org/

Feature	Mini	Mini v2	Huggable
Processor	Intel i7	Intel i7 + Raspberry Pi 3 B+	Raspberry Pi 3 B+
Microcontroller	Arduino MEGA 2560	ESP32	ESP32
Camera	Intel RealSense	-	-
Degrees of freedom	5	5	3
Touch channels	x3	x5	x5
Screens	2	2	-
Microphone	Olympus ME52	Olympus ME52	Olympus ME52
Speaker	Mini stereo speaker	Mini stereo speaker	Mini stereo speaker
Tablet	Samsung Galaxy A12	Samsung Galaxy A12	-
LED indicators	x5	x5	x2
Temperature sensors	x2	x2	x2
Inertial measurement unit	-	-	Adafruit 9-DOF
Battery	-	-	LiPo 3S-1P
Manufactoring cost (euros)	3500	2500	400

Table 5.1: Comparison of the hardware capabilities of the three social robots developed by our research group.

uses nodes as processes that perform computation (functionalities), messages to send asynchronous information between nodes, services to send synchronous information between nodes (clients-service), and topics as channels that connect different nodes throughout which the messages travel. Figure 5.4 shows a general view of our robot architecture, that consist of a memory that stores information about the robot and the user, a *multimodal HRI system*, a *User-adaptive system*, the DMS, and the robots functionalities, called *Skills*.

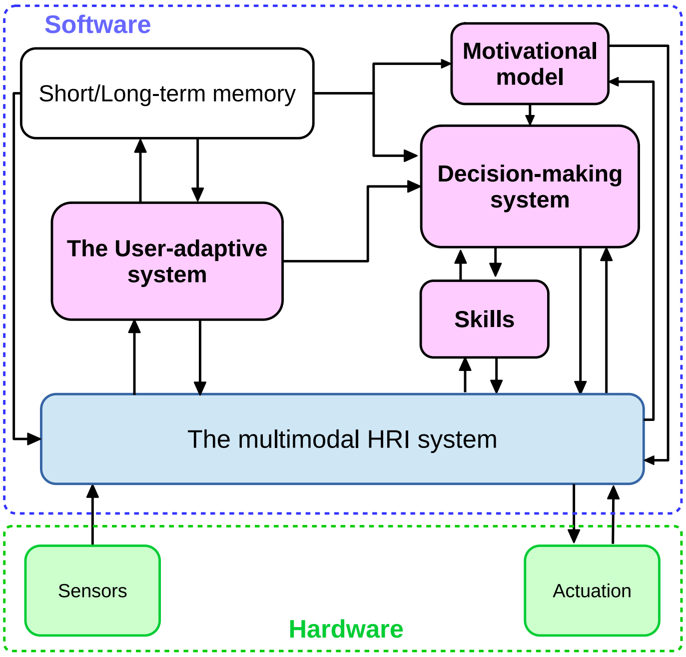


Figure 5.4: General view of the architecture integrated in the social robots of the Robotics Lab research group. In pink, the modules developed in this thesis.

The following sections focus on describing each of the modules shown in Figure 5.4. First, we describe those that have not been developed in this thesis and are essential parts of the architecture. Then, we emphasise (highlighted in pink in Figure 5.4) those that have been designed explicitly during this thesis: the *User-adaptive system*, the DMS (with its motivational model), and some *Skills*. Since the motivational model and the DMS are the cornerstones of this thesis, they respectively have a specific chapter for them in Chapters 6 and 7.

5.2.1. The multimodal HRI system

The bidirectional communication between the robot and people is managed by a multimodal HRI system composed by *Perception manager*, the *HRI manager*, the *Expression manager*, and the *Liveliness manager*. These modules are essential for the architecture's operation since they connect and allow straightforward communication between the sensory and actuation devices of the robot with the rest of the modules for human-robot communication. This system has been fully developed in a parallel thesis of the group [300]. Next, we provide a detailed description of the action of each manager in our HRI system. Figure 5.5 expands the view of the multimodal HRI system introduced in Figure 5.4.

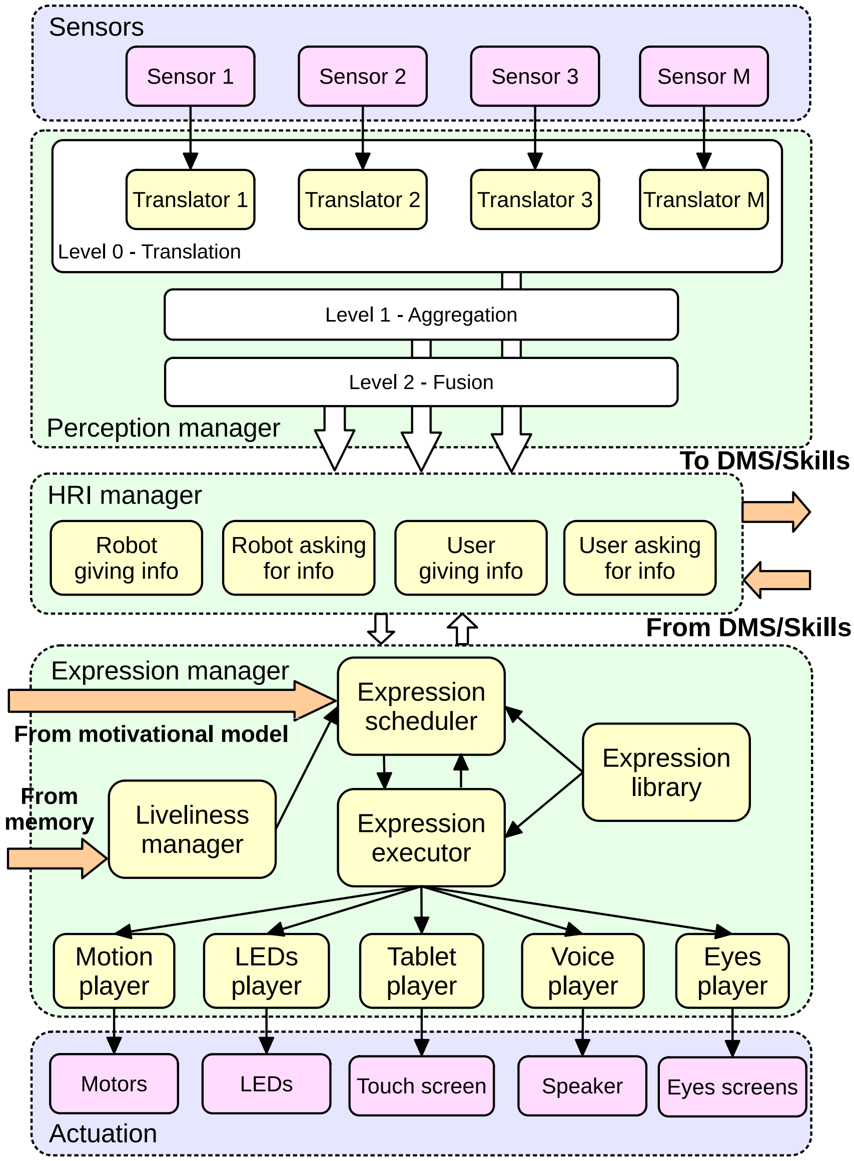


Figure 5.5: The multimodal HRI system divides in the sensors and actuators, the *Perception manager*, the *HRI manager*, and the *Expression manager*. These modules integrates into the software architecture providing a robust interfaces for managing the social communication with the user.

5.2.1.1 Perception manager

The *Perception manager* (see [293]) is a module of our software architecture in charge of integrating the information produced by the sensors of the robot. The primary function of the *Perception manager* is to receive the raw data generated by each sensor yielding a homogenous, aggregated, single message that the other modules of the system can understand. Like Figure 5.5 shows, the *Perception manager* has three abstraction levels.

- The Level 0, called Translation, is formed by a set of dedicated modules that receive the information of a specific sensor or perception unit and translate it into a standard format. This format is an array of key-value pairs, where keys are identifying tags of the information sent and values in the numerical or discrete value of the tag. It is worthy of mentioning that perception units are middle layers between a sensor and the Level 0 that performs data preprocessing (e.g. NLP).

- The Level 1, called Aggregation, receives the information from the Level 0 and combines it producing a single message that contains meaningful perception information from one or more perception units inside a time window of 1 second.
- The Level 2, called Fusion, contains a set of high-level modules that receives information from the two previous levels fusing their information into more complex, high-level messages that fuse information from different sensors. A representative example of this case is the combination of a face detector and a gaze detector to determine whether the robot engages users or, on the contrary, they are not focused on the robot. The goal of the third level is to strengthen the sensory information sent to the other modules of the robot, so future decisions have a more robust basis avoiding possible perception errors.

The different levels of the *Perception manager* directly communicate with the *HRI manager* informing about changes in the environment detected by the sensors of the robot. However, there exist exceptional cases in which other applications can subscribe to the *Perception manager* messages to modify their operation. For example, the *User profiling manager* and the *Context manager* are aware of when a new user approaches the robot, so both can load his/her information to produce a personalised interaction. Similarly, DMS is informed about when the user pretends to interact with the robot using the tablet or calling it by voice commands, letting the decision-making know how to drive the robot's behaviour.

5.2.1.2 Human-robot interaction (HRI) manager

The *HRI manager* [301] is the module that controls the communication between the robot and the user. It is based on an asynchronous bidirectional exchange of information with the *Skills* of the robot using messages called Communicative Act (CA). CAs are communicative units used for requesting or giving information to the user. These units play a pretty important role in abstracting and simplifying the communication process with the user since it provides a standard and unique channel throughout the information with the user flows. Like it is described in [301] the CAs can be classified as robot asks for information, the robot gives information, the user asks for information, and the user gives information, depending on whether the robot or the user initiates the information exchange and the direction of that communication.

The *HRI manager* sequences the interaction with the user but also avoids conflicts in case that more than one *Skill* of the robot require the use of the same devices. Thus, internally, it organises in two levels: the control module and the active CAs. The control module is the sequencer, a kind of planner that manages which CA can be active at each time. Active CAs transmit the information to the *Expression manager*, which decouples the information and sends commands to the actuators. The management of the active CAs is carried out using a priority sorting system with three levels: low, medium, and high. CAs with a higher level of priority are executed first, even if this fact causes an interruption of the current active CA. If two CAs have the same priority, then the one that arrived earlier is deployed first only in case any other CA is currently active. Otherwise, the CA is stored in a queue that passes to the next turn when the current active CA finishes. It is worthy to note that several CAs can be active at the same time, if and only if they use different communication interfaces. Thus, this module is essential in autonomous social robot architectures since it links low-level devices (sensor and actuators) and high-level processes (decision-making).

Like Figure 5.5 shows, the *HRI manager* situates between the *Perception manager* and the *Expression manager*, working in coordination with the DMS to manage perception and control actuation. The execution of certain expressions by the *Expression manager* under the request of the *HRI manager* provides feedback that contains the result of the CA and, in case it requires information from the user, the content of such information. The following section describes the *Expression manager* as the last part of the multimodal interaction modules.

5.2.1.3 Expression manager

The *Expression manager* (see [293]) controls the actuation capabilities of the robot in terms of communication. It receives orders from the *HRI manager* and from the liveliness module about how to control the robot's actuators to produce specific multimodal expressions. On the one hand, the *HRI manager* sends expressions based on practical sensor information. On the other hand, the liveliness module generates spontaneous actuation commands to improve the naturalness of the robots' movements, mainly when the robot does not perform any specific task.

The expressions or gestures received by the *Expression manager* are based on different actuation commands that must be sent separately to each actuator. Thus, the *Expression manager* primary's task is to decouple complex expressions into individual actuation commands and send them to the corresponding actuator at the right time. Considering this fact, the *Expression manager* consists of three levels and a complementary module called *Liveliness manager*, like Figure 5.5 shows.

- The top level is the *Expression scheduler*, a module that receives complex expressions and manages them to define a queue where urgent expressions are top-ranked. The queue is sorted using the priority of each expression, the time it arrived and the interfaces needed to execute it. Thus, the *Expression scheduler* resolves conflicts if more than one expression must be executed simultaneously, needing the same interfaces. It is worthy to note that the *Expression scheduler* also manages the available interfaces and, if possible, coordinates the execution of multiple expressions simultaneously. Once an expression has finished, the *Expression scheduler* verifies if a new expression has to be executed. If a new expression is ready, the *Expression scheduler* sends the expression to the *Expression executor*.

- In the second level, the *Expression executor* is a node that loads the expression information from a predefined library and executes it. The *Expression executor* can manage different expressions acting simultaneously, splitting up the expression into individual actuation commands.

- In the third level, it is possible to find several players. A player is a programme that handles the commands that have to be sent to a specific actuator. Thus, each player is dedicated to an actuator device, receiving the command from the *Expression executor*. Then, the player sends the command received to the actuator manager, which generally is microcontrollers. In our system, there are five players for controlling the robot's motion, the eyes screens, the voice and sounds, the LEDs, and the touch screen. The players can be enabled, disabled, and interrupted when necessary but always controlled by high-level processes.

The expressiveness of the robot requires continuous adaptation depending on the situation where it is involved. Thus, the *Liveliness manager* is included in the software architecture to improve the naturalness of the robot's behaviour and provide a feeling of liveliness to the user. The *Liveliness manager* is a module that generates, using a sine wave that represents the internal rhythms of the robot and the robot's state (e.g. sleeping or playing), appropriate involuntary expressions that enhance its expressiveness.

The sine wave is defined in amplitude and frequency, serving as a robot's autonomic expressiveness modulator. Thus, if the sine waves present a high frequency, the expressions will be executed more often than if the frequency is low. Similarly, the amplitude of the sine function controls the amplitude of the robot's movements. Both parameters depend on the robot's internal state, which can be affected by factors like its physiological deficits, affective state, or the state of the stimuli, among others.

The *Liveliness manager* obtains the information about the robot's circumstances from the robot's memory, where the other nodes of the architecture write the robot's status. On the other hand, the amplitude and frequency of the sine signal that controls spontaneous behaviours directly attack the *Expression manager* on its lower level (players), producing appropriate motor commands for controlling the actuation devices.

5.2.2. The User-adaptive system

In order to achieve good user acceptance and natural interaction, the robot has to adapt its behaviour to the user. Thus, Figure 5.6 expands the view of the *User-adaptive system* presented in Figure 5.4 that integrates various modules for enabling this robot functionality. Next, the *User profiling manager*, the *Context manager*, and the *Scheduler manager* are explained in detail as essential systems required for attaining an adaptive robot behaviour. Next, we provide a detailed description of the function of these managers in the robot architecture.

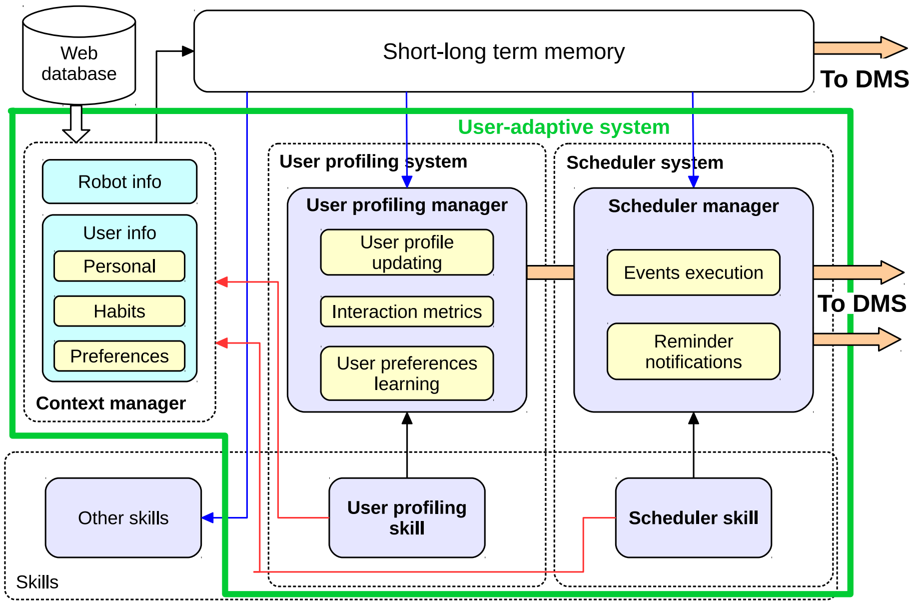


Figure 5.6: The software architecture contains various modules that allow robots to adapt to different users. In this setting, the *Context manager*, the *User profiling manager*, the scheduler system, and the *Skills* interact among them and with the robot's memory to exchange information that allows the DMS to adapt the robot's behaviour to each user. Red arrows show the information that updates the *Context manager* for long-term interactions. Blue arrows indicate the information flow from the robot's memory to the rest of *Managers* and *Skills*. Black arrows indicate the communication of the User profiling and *Scheduler skill*s with their corresponding manager.

5.2.2.1 User profiling system

The User profiling system manages the retrieval of important information of the user to adapt the interaction process. It is made up of two modules: the *User profiling manager* and the *User profiling skill*, like Figure 5.6 shows.

- The *User profiling manager* is a full-time active module that manages how each user interacts with the robot. It receives information from the *HRI manager* about the different actions that the user is taking for generating interaction metrics that shape the quality of the interaction. These metrics are lately used for adapting the robot's behaviour. This module includes the PL system. Using the interaction metrics mentioned previously, it can know how much a user likes each activity.

- The *User profiling skill* is a conditional functionality that only activates when a new user is detected or when the DMS considers it appropriate. The functionality of this module is twofold. First, it informs the *Context manager* (see Section 5.2.2.2) when it has to create a new empty user profile, setting a unique user id to differentiate among different users. Second, this module contains a repertoire of questions to fill the empty user profile by asking the user about personal information (e.g. name, surname, age, sex), preferences (e.g. music, news, games), or habits. As the robot retrieves more and more information about the user, the interaction is more shaped and adequate to him/her.

5.2.2.2 Context manager

The *Context manager* is a module designed to store information that the robot can use to adapt its interaction. It contains information about the robot itself and about the different users that interact with it. Inside the ROS architecture, the information is published in the robot's memory, a kind of virtual board where the nodes of the architecture can efficiently write and read information exchanging it with other nodes.

Principally, the *Context manager* communicates with a web server and the *User profiling manager*. On the one hand, a webserver is an online tool that allows to define basic information about each user like personal data (e.g. name, age, sex) to more complex information like physical or cognitive impairments or proactivity interact. When the robot needs information from a particular user, the *Context manager* downloads the information of the user that it is going to interact with and writes it to the robot's memory to allow other nodes to use it. On the other hand, the *User profiling skill* is a robot functionality that allows the robot to obtain new information. This functionality activates when a never seen user is perceived, indicating to the *Context manager* that it has to create him/her a profile. Then, the user profiling functionality drives the robot to ask the user about different information to build up his/her profile step by step. Figure 5.6 shows how the *Context manager* integrates into the robot's architecture showed in Figure 5.4 and the most important information it contains.

5.2.2.3 Scheduler system

The *Scheduler system* is a software system integrated into our architecture that controls the activities a user has to realise and reminders that the robot has to notify the user. The *Scheduler system* has two modules: the *Scheduler manager* and the *Scheduler skill*, like Figure 5.4 shows.

- The *Scheduler manager* is a full-time running programme that stores the events and reminders of each particular user. The events and reminders are read from the robot's memory. Events are robot activities that a particular user has to execute (e.g. the robot activates a cognitive stimulation therapies at an specific time of the day that has been planned previously) or the robot want to communicate (e.g. inform the user about the last news), and reminders are notifications given to the user (e.g. reminding the user an appointment with the doctor). After processing the events and reminders, the *Scheduler manager* communicates with the DMS when the time of the start of an event or reminder is close.

- The *Scheduler skill* is a functionality that allows users to create, delete, and edit events and reminders by interacting with the robot. The *Scheduler skill* and the *Scheduler manager* communicate between them via a ROS service. Using this service, both nodes exchange information that allows them to manage the events and reminders. An example of this case is the user requesting the robot to remind him in two days an appointment with the hairdresser.

5.2.3. Skills

The robotic platforms presented above can execute many different functionalities called *Skills*. *Skills* were conceived ta flexible and modular applications that can be efficiently included and removed from the robots. In our architecture, the DMS controls the execution of the *Skills*, managing the application that the robot is executing. It is worth mentioning that a *Skill* can be executed in different ways. Each different functionality exhibited by an *Skill* is called *Goal*. The *Goal* indicates how the robot has to execute a specific functionality (e.g. activate the photos skill displaying the monuments goal). The *Skills* devoted to the user's entertainment are classified in categories, subcategories, and goal, as Figure 5.7. This classification simplifies the robot's selection of the user's favourite activities as described in Chapter 9.

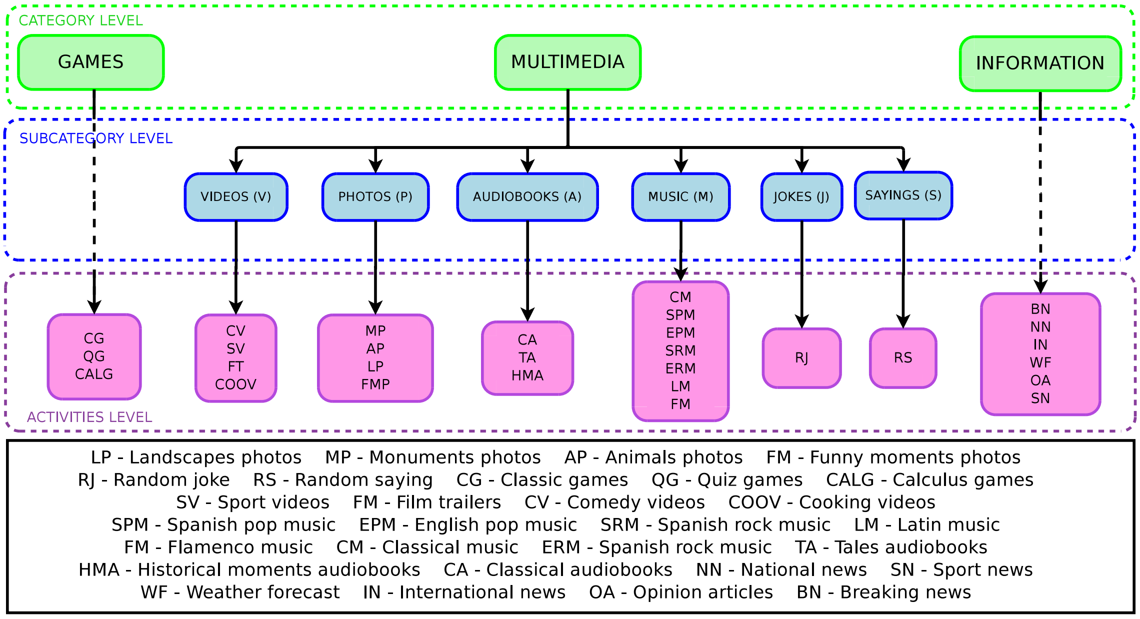


Figure 5.7: Classification of the robot's entertaining activities in categories, subcategories, and goals. This classification is used in the preference learning system to combine proactive autonomous decisions of the robot with giving the initiative to the user allowing him/her to select their favourite activities.

Next, we describe the functionality of each *Skill* integrated into the social robot Mini. The *Skills* specifically designed and developed in this thesis are indicated with an $*$.

- **Sleeping*:** The robot goes to sleep closing its eyes while performing non-verbal sounds like snoring.
- **Waiting*:** The robot relaxes and waits for new upcoming stimuli without doing any specific activity.
- **Communications*:** The robot activates the communications skill to obtain general information from the user during the interaction or to communicate general information about its decisions.
- **Alerts*:** This skill activates when the robot has an important instant communication to the user. For example, the reminders of the *Scheduler manager*.
- **Tablet menus*:** This skill consists of presenting a set of menus on the tablet device that allows the user to select which activity (skill) (s)he prefers to execute from the robot's repertoire.
- **User profiling skill*:** This skill activates when the robot meets a new user or sporadically during the interaction. It contains questions to gather information about the user employed to generate a personalised human-robot interaction.

- **Scheduler skill*:** This skill allows the robot users to create in the robot's agenda new events to execute in the future and reminders that the robot will notify at the specified date and time.
- **Dancing*:** When the dancing skill is activated, the robot plays a song from its repertoire and starts dancing following the rhythm of the song.
- **Quiz game*:** This game consists of different play modes based on answering questions about six different categories: history, science and technology, geography, entertainment, sport, and art and literature.
- **Bingo game:** The skill allows the user to play the bingo game against the robot. The game has different difficulties and modalities that adapt to the user characteristics.
- **Presentation skill:** This skill presents the robot Mini to the audience.
- **Instructions skill:** This skill informs the user about how the social robot Mini has to be used.
- **Cognitive stimulation skill:** The cognitive stimulation package is a complex collection of *Skills* that control the cognitive stimulation therapies designed for assisting older adults with cognitive impairment. The skill allows execution from single exercises of multiple categories like memory, perception, or calculus to complex treatments specifically designed for each patient. This module was part of another PhD thesis of our research group [302].
- **Photos:** The photos skill displays pictures about different topics (e.g. animals, personal photos, or monuments) on the tablet device of the robot.
- **Videos:** The videos skill displays videos on the robot's tablet such as films, cooking recipes, or comedy in the tablet device.
- **Audios:** The audio skill is used for playing sounds like music using the tablet device attached to the robot.
- **News:** The robot looks for information on the internet about a previously selected topic by the user and tells them to him/her.
- **Weather:** The robot looks for the weather forecast on the internet using its current location and tells it to the user.

Chapter 6

Our motivational model

This chapter describes the *Motivational model* proposed in this thesis as a biologically inspired basis for the robot's autonomous behaviour. The model integrates into our software architecture as one of the most important inputs of the DMS, as Figure 6.1.

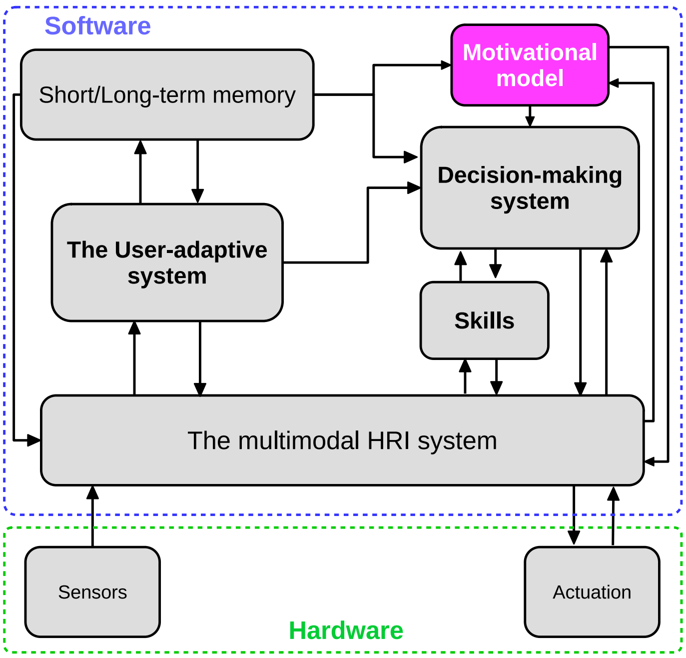


Figure 6.1: Integration of the *Motivational model* into our software architecture.

Its primary function is to emulate processes occurring in the human body, providing a robust mathematical background based on outstanding theories of motivation and affect (previously presented in Chapter 2). Besides, we provide novel modelling equations related to artificial neuroendocrine responses, circadian and ultradian rhythms, autonomic control, and affect generation based on mood and emotion. The following sections explain the modelling in Mini of the processes integrated into our *Motivational model* following the diagram shown in Figure 6.2. First, we present the neuroendocrine responses emulated. Then, we describe how the circadian and ultradian rhythms affect multiple biologically inspired functions. We emphasise the essential modulation of stimuli on neuroendocrine responses, biological processes, and motivated behaviour. Next, we present the somatic and autonomic processes that we emulate in our robot and how they influence its motivational state. Finally, we link motivation and behaviour, indicating how the *Motivational model* informs the DMS about the most appropriate behaviour the robot should perform.

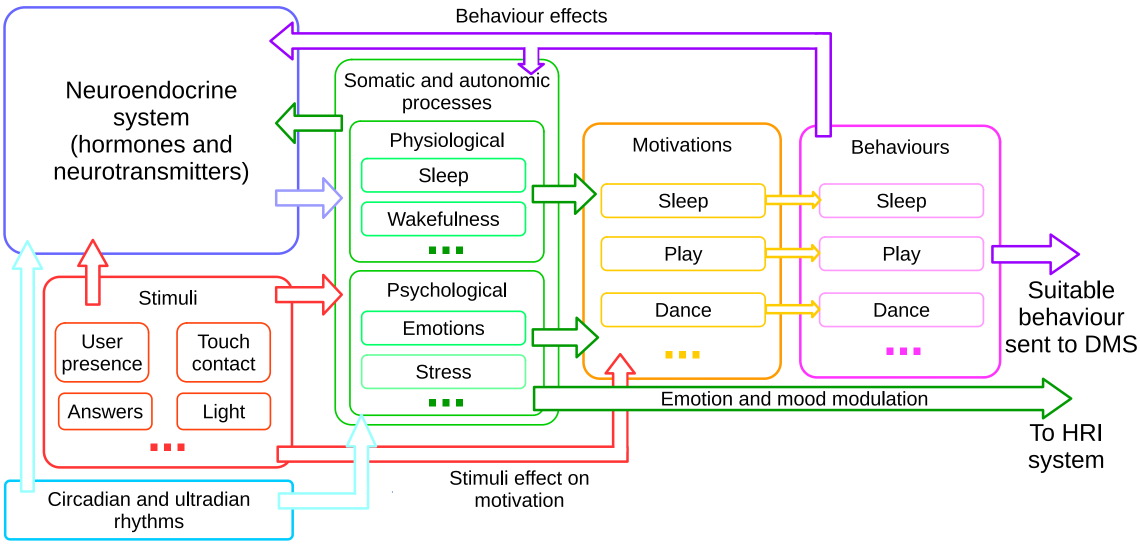


Figure 6.2: General view of our Motivational model.

6.1. Modelling neuroendocrine responses and their interactions

The selection of the most appropriate behaviour by our *Motivational model* originates in the artificial neuroendocrine system presented in this thesis. As Figure 6.2 shows, the neuroendocrine system is a module of the *Motivational model* that emulates neuroendocrine responses in Mini. The neuroendocrine system simulates the levels of artificial hormones and neurotransmitters that control Mini's vital physiological and psychological processes.

Considering the biological basis introduced in Chapter 2, the neuroendocrine system has twelve hormones and neurotransmitters. Each substance regulates one or more physiological and psychological processes in Mini. Thus, Melatonin (MT) regulates the sleep periods of the robot, while Orexin (OX) regulates the activity periods. In our model, Dopamine (DA), Brain norepinephrine (BNE), and Serotonin (SE) primary modulate emotional responses, but also other important autonomic and involuntary processes like heart rate or blinking rate. Oxytocin (OT) and Arginine vasopressin (AVP) control social behaviour, both positive and aggressive interactions. Our model simulates the stress response through the action of Corticotropin-releasing hormone (CRH), Adrenocorticotropin (ACTH), and Cortisol (CT) for managing aversive situations. Finally, we model the effect of Adrenal norepinephrine (ANE) and Epinephrine (EPI) on Mini's autonomic functions like heart rate or pupil size.

In the model, we propose Equation 6.1 to represent the evolution of the neuroendocrine levels. This equation supposes one of the most important contributions in our neuroendocrine system since any previous work in the literature proposes a mathematical equation for describing the daily evolution of hormonal levels. Consequently, and taking strong inspiration from the work developed by Knight [79] characterises the secretion of neuroendocrine substances in humans, Equation 6.1 shows our proposal for modelling the level of hormones and neurotransmitters in our artificial neuroendocrine system. The hormonal level l_i depends on circadian cr_i and ultradian rhythms ur_i, the effect of stimuli se, and the interaction with other substances hi. According to Equation 6.1, each substance can be affected by the intensity se_k of K stimuli and can simultaneously interact hi_p with P substances. The level l_i of a neuroendocrine substance ranges from 0.01 to 1 unit.

$$l_i(t) = cr_i(t) + ur_i(t) + \sum_{k=0}^{K} se_k(t) + \sum_{p=0}^{P} hi_p(t) \tag{6.1}$$

In humans, the number of interactions between neuroendocrine substances is massive [303]. We model neuroendocrine interactions as paired effects that implicate a source and target substances.Figure 6.3 shows the neuroendocrine interactions included in our model. These interactions have been obtained from the literature presented in Chapter 2 to simulate in our system essential biological interactions occurring in humans.

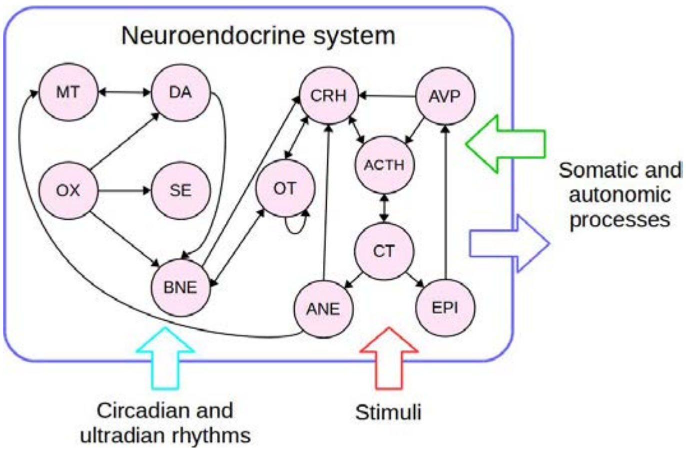


Figure 6.3: Neuroendocrine interactions represented in our model.

On the one hand, the target substance receives the effect, altering its levels. On the other hand, the source substance is the substance that carries out the effect. In our system, the effect can be stimulatory or inhibitory. Stimulatory (S) interactions increase the level of the target substance, while inhibitory (I) interactions decrease the level of the target substance. Like Equation 6.2 shows, we propose that each neuroendocrine interaction hi_p is calculated considering an empirical weight δ_b ranging $[-1, +1] - \{0\}$ that modulates the type and the intensity of the interaction and the level of the source substance l_{source} in a specific time step t.

$$hi_{\text{p source}}(t) = \delta_b \cdot l_{source}(t) \tag{6.2}$$

Table 6.1 shows the neuroendocrine interactions included in our *Motivational model*. To simplify the modelling process, we have defined three typical values for the δ_b weight: ± 0.001 if the interaction intensity is low, ± 0.05 if the interaction intensity is moderate, and ± 0.01 if the interaction intensity is strong. The sign of δ_b indicates if the type is stimulatory (+) or inhibitory (−). It is worth mentioning that some δ_b values in Table 6.1 have been carefully adjusted and do not use the predefined weights since they represent an important function in the robot's internal state. An example of these effects are the relations between the hormones CRH, arginine vasopressin, ACTH, cortisol, adrenal norepinephrine, and epinephrine in the HPA axis.

Source substance	Target substance	Type	δ_b
Melatonin	Dopamine	I	−0.01
Orexin	Brain norepinephrine	S	+0.05
	Dopamine	S	+0.05
	Serotonin	S	+0.05
Dopamine	Melatonin	I	−0.01
	Brain norepinephrine	S	+0.01
Oxytocin	Oxytocin	S	+0.01
	CRH	I	−0.01
	Brain norepinephrine	I	−0.01
Brain norepinephrine	CRH	I	−0.01
Arginine vasopressin	CRH	S	+0.6
	ACTH	S	+0.1
CRH	ACTH	S	+0.9
	Oxytocin	S	+0.01
	Brain norepinephrine	S	+0.01
	CRH	I	−0.005
ACTH	Cortisol	S	+0.8
	CRH	I	−0.01
Cortisol	Epinephrine	S	+0.4
	Adrenal norepinephrine	S	+0.1
	CRH	I	−0.005
	ACTH	I	−0.005
Adrenal norepinephrine	CRH	S	+0.01
	Melatonin	I	−0.01
Epinephrine	CRH	S	+0.01
	Arginine vasopressin	S	+0.01

Table 6.1: Existing neuroendocrine interactions in the artificial neuroendocrine system of the robot Mini.

6.2. Modelling the influence of stimuli

The stimuli we perceive from the environment and how we cognitively interpret them influence many diverse processes in our organism [304]. Like Figure 6.2 shows, in our *Motivational model*, stimuli influence Mini's internal state at three different levels.

- **Neuroendocrine system:** Stimuli modify the levels of neuroendocrine substances by stimulating or inhibiting their secretion. We model the influence of an stimuli se_k depending on its intensity si_k, a weight α_k that represents the physical impact of the stimuli, and a weight β_k that represents how the agent cognitively interprets the stimuli. The intensity of an stimulus is provided by the perception system of the robot ranging $[0, 100]$, the weight α_k ranges $[-1, +1] - \{0\}$, and the weight β_k ranges $[-2, +2] - \{0\}$. This modelling translates into Equation 6.3, a mathematical formula proposed in this thesis.

$$se_k(t) = \alpha_k \cdot \beta_k \cdot si_i(t) \tag{6.3}$$

In our model, all weights α_k are empirical and fixed values since they represent physiological effects. However, in some particular cases, the cognitive interpretation that the weight β_k represents varies based on the value of a biological process, describing an allostatically regulated process [73]. An example of these regulatory mechanisms is the processes related to socialisation. In our model, Mini's social behaviour varies depending on the user that it is facing. If the user is friendly, Mini will promote social behaviour, while if the user is aversive, Mini will refuse the interaction. This allostatic mechanism is deeply explained in Section 11.2, defining how the social behaviour of the robot varies depending on the pair-bonding between the Mini and each person that interacts with it.

- **Somatic and autonomic processes:** The effect of stimuli on biological processes is modelled using Equation 6.3, as we consider that the stimuli effect on somatic and autonomic processes is akin to their influence on neuroendocrine substances [79].

- **Motivation:** Stimuli are important behavioural triggers [202]. In specific circumstances, the perception of a stimulus defines our behaviour even if we do not have any specific necessity in our organisms (e.g. eat palatable food when we are not hungry at all). Thus, we include in our model the influence of stimuli on motivation as an important driver of the robot's behaviour. The modelling of the influence of stimuli on motivational states is addressed in Section 6.6.

Drawing on Tomkins stimuli classification [192], Mini can perceive two kinds of stimuli: sensorial and cognitive. On the one hand, sensorial (S) stimuli are perceived by the environment. On the other hand, cognitive (C) stimuli require an interpretation of a situation. Currently, the social robot Mini can perceive the following stimuli:

- **Light level:** The robot can perceive the illumination level using a photoreceptor placed on its head.

- **Ambient noise:** Mini can perceive the noise in the ambient using the microphone integrated into its body.

- **Hits:** Mini can perceive when the user hits it using the capacitive sensors placed on the belly, right shoulder, and left shoulder. We consider hits as tactile contacts that have a duration below one second.

- **Strokes:** Similarly to the strokes, the robot can perceive when the user is stroking it. We consider strokes as tactile contacts with a duration above one second.

- **Correct and wrong answers:** Mini can play different games with the user. During these games, the user can correctly or incorrectly answer the questions of the robot.

- **Mini fulfils or not fulfils a goal:** Mini has specific tasks to accomplish. The fulfilment or not of the robot's task suppose a cognitive stimulus for the robot.

- **Friendly and aversive users:** Depending on how the user behaves with Mini, it can cognitively interpret if the user is friendly or an enemy depending on the PB between them. The robot's behaviour may vary depending on this stimulus, adapting to the situation that the robot is living.

Table 6.2 shows the stimuli effects on Mini's neuroendocrine substances, somatic, and autonomic processes. A stimulus can have more than one effect, each effect relying on a physical modulation value α_k and a cognitive modulation value β_k. Both parameters have been empirically set to yield the desired behaviour in Mini. Depending on the sign of α_k and β_k, the impact can be stimulatory (+) or inhibitory (−). In order to simplify the modelling process, in most cases, the value of both parameters will be ±0.001 if the impact of the stimulus is soft, ±0.005 if the impact is moderate, ±0.01 if the impact is strong, and 1 if it does not have any impact. The exceptions are the effect of light on melatonin and the pupil size and the effect of strokes on the social need, which have been precisely adjusted to follow a specific evolution. As Table 6.2 shows, the weight β_k is only used in two effects that represent the modulation of oxytocin release when a friendly user is perceived or when the robot is stroked. The modulation of β_k depends on the human-robot Pair-bonding (PB) value that reflects the relationship between the robot and the user.

Stimulus	Type	Neuroendocrine substance	Somatic/autonomic process	α_k	β_k
Light	S	Melatonin	-	-0.008	1
		Orexin	-	$+0.005$	1
		-	Pupil size	$(-6/100+2)$	1
Ambient Noise	S	CRH	-	$+0.005$	1
Hit	S	Arginine Vasopressin	-	$+0.05$	1
		CRH	-	$+0.001$	1
		Brain norepinephrine	-	$+0.01$	1
Stroke	S	Oxytocin	-	$+0.05$	$+0.02$*PB
		Dopamine	-	$+0.001$	1
		Brain norepinephrine	-	-0.01	1
		-	Social need	-0.002	1
		Serotonin	-	$+0.01$	1
Correct Answer	C	Dopamine	-	$+0.01$	1
		Serotonin	-	$+0.01$	1
		Brain Norepinephrine	-	-0.005	1
Wrong Answer	C	Dopamine	-	-0.005	1
		Serotonin	-	-0.005	1
		Brain Norepinephrine	-	-0.005	1
Mini fulfils a goal	C	Dopamine	-	$+0.01$	1
		Serotonin	-	$+0.01$	1
Mini not fulfils a goal	C	Dopamine	-	-0.005	1
		Serotonin	-	-0.005	1
		Brain Norepinephrine	-	-0.005	1
Friendly User Presence	S	Dopamine	-	$+0.005$	1
		Oxytocin	-	$+0.005$	$+0.02$*PB
		CRH	-	$+0.001$	1
		Brain Norepinephrine	-	$+0.005$	1
		Serotonin	-	$+0.005$	1
Aversive User Presence	S	Dopamine	-	$+0.005$	1
		Oxytocin	-	$+0.005$	1
		Serotonin	-	$+0.005$	1

Table 6.2: Stimuli with influence on the artificial endocrine system, somatic, and autonomic processes of Mini.

6.3. Modelling circadian rhythms

Numerous studies demonstrate that the human body presents periodic variations in critical biological functions [103, 130]). When the periodicity of the process is close to a natural day (24h), the variation is denoted as circadian rhythm. Previous works in mathematical modelling of circadian rhythms propose cosine, and sigmoid functions as precise approaches of these biological oscillations [305, 306]. Following these definitions, we opted for approximating the circadian rhythms exhibited by the neuroendocrine substances and autonomic processes in our *Motivational model* using a cosine wave (Equation 6.4) and the difference of two sigmoid functions (Equation 6.5). The selection of the circadian function for each biological function was carried out considering neuroendocrine studies that define the evolution of the process, like Table 6.3 shows.

$$cr(t) = br + ar \cdot \cos(2\pi \cdot (\frac{t - \mathrm{Tz}}{24})) + n(t) \tag{6.4}$$

$$cr(t) = br + ar \cdot \left| \frac{1}{1 + e^{-(t-\mathrm{Tz})}} - \frac{1}{1 + e^{-dr(t-\mathrm{Tzd})}}) \right| + n(t) \tag{6.5}$$

In the previous equations, t is the time of the day in hours, br is the basal level of the process, ar is the amplitude of the signal, Tz is the time of the day where the signal reaches the secretion peak in hours, Tzd is the time of the day where the sigmoid evolution decreases to basal levels, and $n(t)$ is a predefined basal value from 0.001 to 0.002 that is randomly calculated in every time step and used to avoid a value of 0 units in the process.

Table 6.3 shows the circadian rhythm associated to the neuroendocrine substances included in our *Motivational model*. The parameters of each function have been thoroughly adjusted to attain the desired behaviour of the robot.

Substance	**Circadian function**	**References**
Melatonin	$0.6 + 0.29 \cdot \cos(2\pi \cdot (\frac{t-1}{24})) + n(t)$	[82]
Orexin	$0.3 + 0.3 \cdot \cos(2\pi \cdot (\frac{t-20}{24})) + n(t)$	[93, 307]
Dopamine	$0.05 + 0.1 \cdot \left|(\frac{1}{1+e^{-(t-8)}}) - (\frac{1}{1+e^{-0.3(t-20)}})\right|) + n(t)$	[308]
Serotonin	$0.1 + 0.05 \cdot \cos(2\pi \cdot (\frac{t-18}{24})) + n(t)$	[109]
Brain Norepinephrine	$0.05 + 0.1 \cdot \left|(\frac{1}{1+e^{-(t-8)}}) - (\frac{1}{1+e^{-0.3(t-17)}})\right| + n(t)$	[309]
Oxytocin	$0.05 + 0.03 \cdot \cos(2\pi \cdot (\frac{t-10}{24}))) + n(t)$	[310]
Arginine vasopressin	$0.05 + 0.03 \cdot \cos(2\pi \cdot (\frac{t-11}{24})) + n(t)$	[311]
CRH	$0.05 + 0.03 \cdot \cos(2\pi \cdot (\frac{t-6}{24}))) + n(t)$	[130]
ACTH	$0.05 + 0.01 \cdot \cos(2\pi \cdot (\frac{t-7}{24})) + n(t)$	[130]
Cortisol	$0.05 + 0.01 \cdot \cos(2\pi \cdot (\frac{t-8}{24}))) + n(t)$	[130]
Adrenal norepinephrine	$0.05 + 0.01 \cdot \cos(2\pi \cdot (\frac{t-10}{24})) + n(t)$	[130]
Adrenal epinephrine	$0.05 + 0.01 \cdot \cos(2\pi \cdot (\frac{t-12}{24})) + n(t)$	[130, 143]

Table 6.3: Circadian evolution of the substances available in the artificial endocrine system of the social robot Mini.

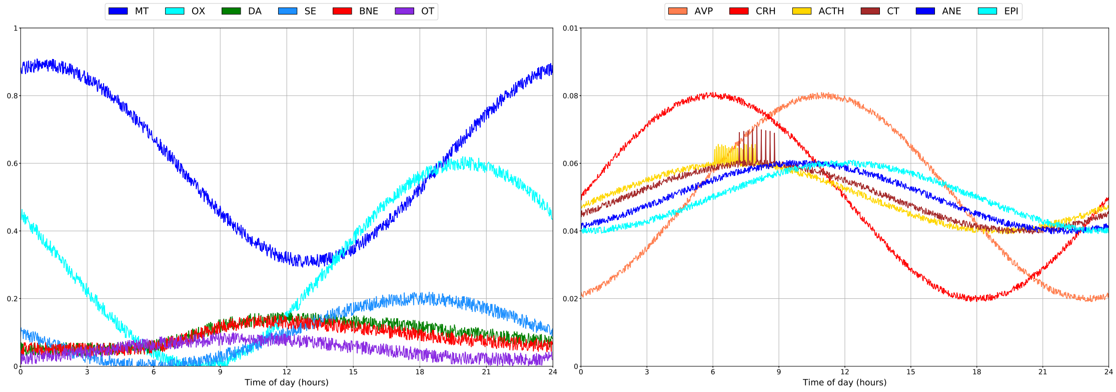


Figure 6.4: Circadian and ultradian rhythms of the twelve neuroendocrine responses included in our model.

Next, in Section 6.5, we define the circadian rhythms exhibited by the somatic and autonomic processes included in the *Motivational model* integrated into our social robot Mini.

6.4. Modelling ultradian rhythms

In humans, some biological functions experience short pulsatile periodic variations in specific moments of the day [151]. These variations that typically have a period smaller than a natural day (24h) are called ultradian rhythms. We model ultradian rhythms as episodic risings in the levels of specific substances during some hours of the day. Thus, ultradian rhythms are characterised by a starting time in hours, an ending time in hours, a period in seconds representing how often the increment occurs, and the numerical value representing the increment itself ur. Table 6.4 shows the ultradian rhythms included in our *Motivational model* and their features.

Substance	**Start time**	**End time**	**Period(s)**	**Effect**	**Reference**
Cortisol	7am	9am	800	0.01	[130]
ACTH	6am	8am	200	0.005	[130]

Table 6.4: Ultradian evolution in the artificial neuroendocrine system of the social robot Mini.

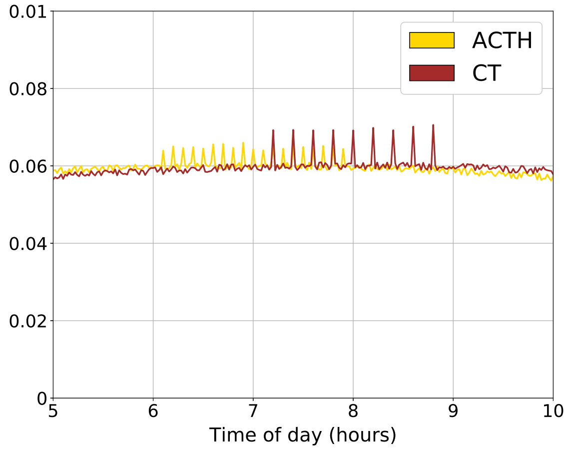


Figure 6.5: ACTH and cortisol ultradian rhythms in the early morning. Cortisol risings range from 7-9 from 7am to 9am. ACTH ultradian rhythm is more repetitive, but with less intensity.

6.5. Modelling somatic and autonomic processes

Human bodies are complex systems that follow precise internal regulations to maintain all their functions in good condition [75]. The biological processes that take place in our organism are divided into those controlling physiological functions like heart rate or sleeping and those regulating our psychological and cognitive development like emotion or social relationships [304]. These biological functions are strong motivators of our voluntary and involuntary actions, exerting a considerable influence on how we behave.

Depending on the kind of behaviour they elicit, it is possible two find two kinds of processes [303]:

- **Somatic processes:** Affect motivational states related to the execution of voluntary behaviour like sleeping or playing. These processes are affected by stimuli and neuroendocrine substances.
- **Autonomic processes:** Directly control involuntary behaviours of our organisms like heart rate, blinking, or pupil dilation. These processes are controlled by endocrine glands or organs, exhibiting precise circadian rhythms and being affected by environmental stimuli and neuroendocrine substances.

In our *Motivational model*, we consider those processes that control autonomic functions (autonomic processes) and those that control our behaviour throughout our motivational state (somatic processes). It is worthy of noting that, besides, biological processes can be physiological (PHY) or psychological (PSY).

We model somatic processes as homeostatically controlled variables that evolve with time. As Figure 6.2 shows, these processes are influenced by neuroendocrine substances, stimuli, and the interactions among them. Since each somatic function controls a specific process in Mini, we opted for defining a unique equation for modelling the evolution of each somatic process. The formula that defines how each somatic process evolves has been empirically adjusted considering the behaviour we want our robot to exhibit and using neuroendocrine studies like for example, the equations for emotion activation, that use the levels of dopamine, serotonin, and brain norepinephrine following the Lövheims's definition [191]. Nevertheless, all the somatic processes in our model range [0, 100] units presenting an initial value *iv* from which they start evolving and an ideal value *iv* where the state of the process is optimal for the agent. As we mentioned before, somatic processes are important behavioural triggers influencing our motivations. More specifically, somatic processes urge us to take specific actions when they present deficits. We model the deficit of a somatic process as the absolute difference of its current value *cv* and its ideal value *iv* [312], like Equation 6.6 expresses. As we present in Section 6.6, motivation lead to behaviour influenced by somatic deficits.

$$d_i = |cv - iv| \tag{6.6}$$

Table 6.5 shows the somatic processes included in the *Motivational model* of the social robot Mini. The calculation of the value that each somatic process presents depends, in some cases, on the value presented in the previous interaction (PV). Moreover, some processes interact, creating a complex network of interactions.

Biological process	Type	Initial value	Ideal value	Activation
Sleep	PHY	0	0	$\mathrm{PV} + 0.002 * \mathrm{MT}$
Wakefulness	PHY	50	50	$100 * \mathrm{OX}$
Social need	PHY	0	0	$\mathrm{PV} + 0.02 * \mathrm{OT} - 0.02 *$ Stroke intensity
Entertainment	PSY	0	0	$\mathrm{PV} + 0.02 * (\mathrm{DA} + \mathrm{BNE})/2$
Pleasantness	PSY	100	100	All deficits
Arousal	PSY	50	50	(Stress + Wakefulness)/2
Joy	PSY	0	0	$100 * \mathrm{DA}^2 * \mathrm{SE}^2 * (1 - \mathrm{BNE})^2$
Sadness	PSY	0	0	$100 * (1 - \mathrm{DA})^2 * (1 - \mathrm{SE})^2 * (1 - \mathrm{BNE})^2$
Anger	PSY	0	0	$100 * \mathrm{DA}^2 * \mathrm{BNE}^2 * (1 - \mathrm{SE})^2$
Surprise	PSY	0	0	$100 * \mathrm{BNE}^2 * \mathrm{SE}^2 * (1 - \mathrm{DA})^2$
Pair-bonding (PB)	PSY	50	100	$\mathrm{PV} + 0.2 * \mathrm{OT} - 0.2 * \mathrm{AVP} - 0.01$
User rejection	PSY	50	0	100 - PB
Stress	Both	0	0	$100 * \mathrm{CT}$

Table 6.5: Definition of the somatic processes included in the Motivational model. PV: Previous value of the process, MT: Melatonin, OX: Orexin, OT: Oxytocin, DA: Dopamine, BNE: Brain norepinephrine, SE: Serotonin, AVP: Arginine vasopressin, CT: Cortisol, PB: Pair-bonding.

On the other hand, we model the autonomic processes to control involuntary processes like Mini's heart rate, pupil size, or breathing rate. Unlike the somatic processes defined in Table 6.5, which all range [0, 100], each autonomic process has a range obtained from neuroscientific studies. In the model, all autonomic processes have a circadian rhythm that defines the basal evolution of the process. Then, the neuroendocrine substances CRH, dopamine, serotonin, epinephrine and norepinephrine modulate their evolution during the day, defining the reactive effect to, for example, unexpected situations. Like the somatic processes, the calculation of the value that each autonomic process presents in a particular time step t depends on a specific equation empirically adjusted to allow Mini to exhibit the behaviour we desire. Note that all autonomic processes modulate a physical variable in Mini, as we show in Chapter 11.

Table 6.6 shows the autonomic processes included in the architecture for controlling Mini's heart rate, breathing rate, blink rate, locomotion activity, and pupil size. The physical function of these variables are:

- **Heart rate:** Variable that controls the simulation of the beating expressed by a coloured RGB LED placed in Mini's chest.
- **Breathing rate:** Variable that controls the frequency and volume with which a subtle sound is simulating Mini's breath is played.
- **Blink rate:** This variable controls the frequency with which the animated screens emulate Mini's eyes blink.
- **Locomotor activity:** Variable that controls the frequency and amplitude of the movements exhibited by Mini.
- **Pupil size:** Variable that regulates the pupil size of the animated eyes of Mini.

Autonomic process	Range	Circadian rhythm	Activation	References
Heart rate	$[50, 200]$ beats per minute	$60 + 3 \cdot \cos(2\pi \cdot (\frac{t-12}{24}))$	$100 * \frac{(EPI+ANE)}{2}$	[144, 313] [314]
Breathing rate	$[10, 30]$ breaths per minute	$20 + 5 \cdot \cos(2\pi \cdot (\frac{t-7}{24}))$	$5 * CRH - 5 * DA$	[315] [316] [317]
Blink rate	$[0, 40]$ blinks per minute	$20 + 10 \cdot \cos(2\pi \cdot (\frac{t-22}{24}))$	$10 * DA * SE$	[318] [319]
Locomotor activation	$[0, 100]$ units	$50 + 1 \cdot \cos(2\pi \cdot (\frac{t-7}{24}))$	$100 * \frac{(DA+SE)}{2}$	[144] [320]
Pupil size	$[2, 8]$ mm	None	$ANE + 3 * EPI$	[314, 321]

Table 6.6: Definition of the evolution of the autonomic physiological processes that control the heart rate, breathing rate, blink rate, locomotion, and pupil size. EPI: Epinephrine, ANE: Adrenal norepinephrine, CRH: Corticotropin realising-hormone, DA: Dopamine, SE: Serotonin.

6.6. Modelling motivation

Motivations are psychological states derived from our internal and external situation [202]. Thus, motivational states reflect the agent's needs, intentions, perceptions, and desires, translating them into specific behaviour. As Table 6.8 and Figure 6.2 shows, in our model, each motivational state activates an specific behaviour. Considering important motivational theories [197, 201, 202, 312, 322], we model the robot's motivational states as psychological processes dependent on the value of biological processes, the deficits of biological processes, and the intensity with which we perceive environmental stimuli. Consequently, in this thesis, we propose a new Equation 6.7 for characterising the intensity of Mini's motivational states. Each motivational state has an intensity denoted by m_i that, like Equation 6.7 shows, depends on the value of the biological processes with an impact on the motivation $bp_b(t)$, the deficits of the biological processes with an impact on the motivation $d_p(t)$, and the intensity of some stimuli with an impact on the motivation $si_s(t)$. As we later describe, active motivations compete to become dominant and rule the robot's behaviour.

$$m_i(t) = \prod_{p=0}^{P} d_p(t) * \prod_{b=0}^{B} bp_b(t) * \prod_{s=0}^{S} si_s(t) \ \forall \ d_p, bp_b, si_s \ \in \ m_i \tag{6.7}$$

The Equation 6.7 we propose derives from Equation 6.8, a well-known formula for computing motivational intensities [199, 268]. Equation 6.8 indicates that the intensity of a motivational state in time step t depends on a biological deficit d_i plus the value of the deficit modulated by the intensity of external stimuli (si_i). Although Equation 6.8 can be extended to situations where the motivation depends on more than one biological process and stimulus, from our point of view, it does not provide enough flexibility since it can not explain purely reactive motivations that do not require a deficit in the agent for becoming active.

$$m_i(t) = d_i(t) + d_i(t) \cdot si_i(t) \tag{6.8}$$

Because of the previous definitions, Table 6.7 shows Mini's motivational states and the equation that activates each of them. In our model, we define two kinds of motivations: proactive or reactive [203]. Proactive (P) motivational states produce a voluntary reaction in the agent, while reactive (R) produce involuntary responses to unexpected or undesired environmental changes (stimuli). In most cases, the behaviour elicited by a reactive motivation takes a few seconds (e.g. reacting to extreme ambient noise when playing a game).

Motivation	Type	Threshold	Activation
Sleep	P	30	Sleep df ∗ (100 − Wakefulness)
Awake	P	5	Wakefulness
Play alone	P	30	Entertainment df ∗ Wakefulness ∗ (100 − User presence)
Play with the user	P	30	Entertainment df ∗ Wakefulness ∗ User presence
Socialise	P	30	Social df ∗ Wakefulness ∗ User presence
Affect	P	30	Social df ∗ Wakefulness ∗ (100 − User presence)
Relax	P	30	Stress df ∗ Wakefulness
React to User presence	R	100	↑ User presence
React to Correct Answer	R	100	↑ Correct answer
React to Wrong Answer	R	100	↑ Wrong answer
React to Hit	R	40	Hit intensity
React to Stroke	R	40	Stroke intensity

Table 6.7: List of motivations of the robot and the equation defining their intensity. Some motivations are reactive, leading to punctual responses to certain stimuli. Reactive motivations (R) are fully dependent on abrupt changes in the perception of some stimuli (marked with ↑ or ↓). Conversely, proactive motivations (P) usually drive voluntary behaviour, leading to behaviours with long-lasting effects. Motivations with a level of intensity below its threshold can not be a candidate to be dominant. df indicates the deficit of the associated process.

The behaviour recommended by the *Motivational model* to the DMS depends on the robot's motivational state. Our model defines the motivational state as the motivation with the highest level of intensity among all active motivations. Thus, drawing on previous motivational models [198, 323, 324], we use a winner-take-all approach to compute the dominant motivation in every time step. To avoid those motivations with shallow levels becoming active, each motivation has a threshold value that has to be overcome for being a candidate to become dominant. If any motivation is above its threshold, the Awake motivation will be dominant, maintaining the robot active.

6.7. Mini's behaviour

The *Motivational model* we proposed in this thesis aims at providing the DMS the most appropriate behaviour according to the artificial physiological and psychological state of the robot. Table 6.8 shows the list of behaviours that the social robot Mini can execute. These behaviours have been selected considering the entertainment and companionship applications of Mini. It is important to note that the behaviour suggested by the *Motivational model* to the DMS is not always executed because the DMS considers in the decision-making process other information such as the user's agenda or petitions.

The activation of a particular behaviour depends on the robot's dominant motivation, which defines the most important processes the robot needs to pay attention to most urgently. Thus, each behaviour is directly tied to one motivation. Every time the dominant motivation changes, the *Motivational model* updates the suggested behaviour to the DMS. Like Table 6.8 shows, the execution of behaviour affects the internal biological processes of Mini, principally seeking to reduce physiological and psychological needs. Besides, in some circumstances, the execution of behaviour requires the presence or perception of a particular stimulus since stimuli are behavioural triggers [203]. It is worth mentioning that the *Play* behaviour clusters a broad range of entertainment activities whose selection depends on the Preference Learning (PL) system previously described in Chapter 4.

Behaviour	**Related motivation**	**Effect on substances and processes**	**Requisites**
Sleep	Sleep	Reduces sleep need	None
Stay awake	Awake	None	None
Dance	Play alone	Reduces entertainment need	None
Play	Play with user	Reduces entertainment need Reduces social need	User present
Request affect	Affect	None	None
Talk with people	Socialise	Reduces social need	User present
Meditate	Relax	Reduces CRH levels	None
Welcome the user	React to User Presence	None	User present
Congratulate	React to Correct Answer	None	User present
Encourage the user	React to Wrong Answer	None	User present
Complain about hit	React to Hit	None	User present
Thank the user	React to Stroke	None	User present

Table 6.8: List of behaviours in the social robot Mini.

6.8. Modulating the robot's expressiveness with emotion and mood

Our affective state not only defines our decision-making but also modulates our expressiveness [179]. In our model, like Figure 6.2 shows, emotions are important regulators of our motivational state, and both mood and emotion regulate the robot's expressiveness. While we consider emotion intense emotional states that rule Mini's expressiveness in the short term, mood rules our affective state and expressiveness in the long run [195].

On the one hand, like we previously presented in Section 6.5, our model considers four primary emotions (anger, joy, sadness, and surprise) as psychological processes with influence on behaviour using the Lövheim's monoamine model [191]. The intensity of these emotions is calculated considering the levels of the monoamines dopamine, serotonin, and brain norepinephrine, as stated by Lövheim. In summary, the robot can express four emotional states that primarily arise due to unexpected stimuli and specific situations.

On the other hand, drawing on previous multi-dimensional theories of mood generation [175, 190, 195], we model mood states using the degree of pleasantness and arousal of the agent. On the one hand, we propose that the pleasantness value depends on the deficits of the agent, as Equation 6.9 shows. On the other hand, we propose that the arousal value that determines its activation level depends on the levels of wakefulness and stress, like Equation 6.10 shows. The pleasantness and arousal values range [0, 100], defining a bidimensional valence-arousal space where different moods are located.

$$\text{pleasantness} = \sum_{i=0}^{N} d_i \tag{6.9}$$

$$\text{arousal} = \frac{\text{Stress} + \text{Wakefulness}}{2} \tag{6.10}$$

Figure 6.6 shows the position in the bidimensional space of the moods included in our *Motivational model*. Mini's mood can be neutral, happy, excited, angry, afraid, sad, depressed, calmed, and contented, drawing on Zhang et al. [195] study.

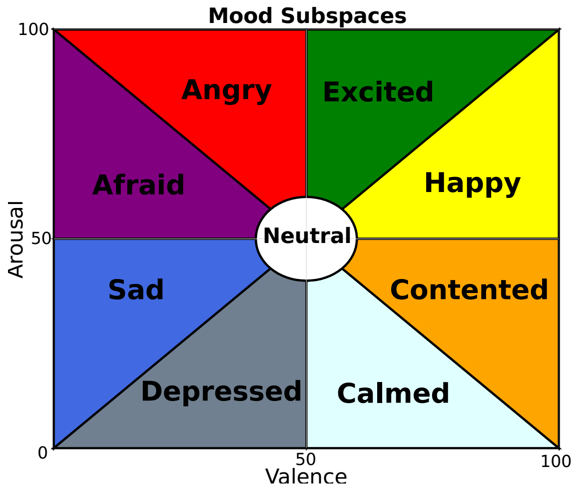


Figure 6.6: Pleasantness-arousal bidimensional space used for determining the robot's mood.

Emotion and mood modulate the robot's expressiveness, allowing it to convey its affective state. As Figure 6.2 shows, the *HRI system* manages the robot's expressiveness using the information sent by the *Motivational model*. The generation of the suitable expression of the affective state is further described in [300].

6.9. Conclusion

In this chapter, we have presented our *Motivational model* for the autonomous behaviour of the social robot Mini using biologically inspired processes. This system aims to provide the DMS with information of the robot's internal state, endowing it with the possibility to behave according to its emulated artificial deficits and react to unexpected situations. Besides, our model simulates affective processes considering emotion and mood, attempting to emulate how affective phenomena influence human motivated behaviour and expressiveness.

In the following chapter, we continue describing our architecture, focusing on how our DMS uses the information provided by the *Motivational model* and other modules included in the software architecture to decide the best action Mini has to execute after the evolution of its internal and external circumstances.

Chapter 7

Our decision-making system

This chapter describes the autonomous Decision-Making System (DMS) developed in this thesis. First, we provide a general overview of the system's requirements. Then, we enumerate the steps followed by the decision-making loop to make a decision every time step. Finally, the chapter explains the system's conceptual design, ending with its integration into the social robot Mini software architecture. Our DMS integrates into our software architecture as Figure 7.1 shows, endowing the robot with autonomous behaviour.

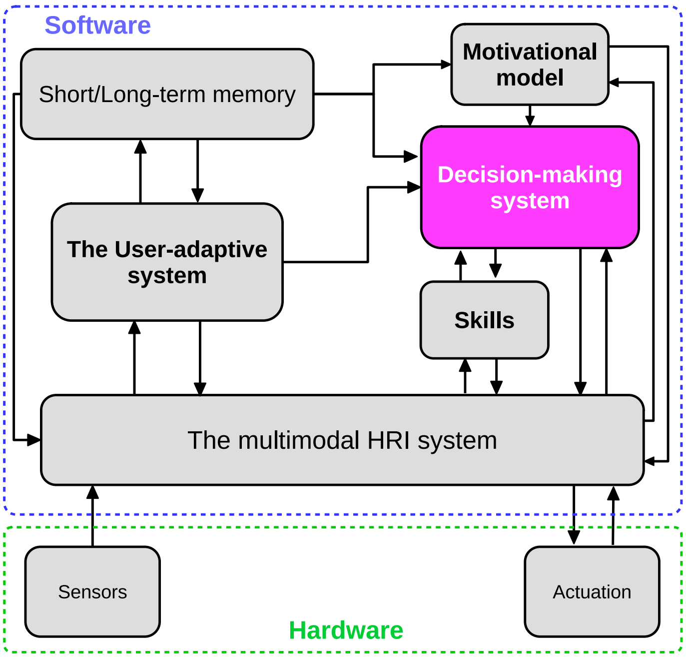


Figure 7.1: Integration of the DMS into our software architecture.

7.1. General overview

The DMS developed in this thesis aims at allowing our social robots to deploy an autonomous and natural behaviour for long-lasting interactions. The decision of the most appropriate behaviour in every moment depends on many internal and external circumstances that define the robot's state and the user's needs and preferences. In the design phase, we defined the requisites of our DMS and its principal functionalities. Next, we list the principal features we wanted to include in our system.

- The priority of the robot is always the user. Mini has to orient its actions to the user petitions. We assumed that Mini could face very proactive users who prefer to select on their own which activity to execute and not proactive users who require the robot to suggest to them the activity to execute. The proactivity level is a numerical value from 0 to 5 that indicates Mini how much proactive the user is. This parameter is stored in the robot's *Memory*. The solution we propose to this problem is to combine the robot making autonomous decisions with allowing users to decide which activity to perform. If the user proactivity is high, the robot is more likely to cede the initiative to select their favourite activity. Otherwise, if the user's proactivity is low, the robot is more likely to make autonomous decisions.

- When the robot is not interacting with any user, it is important to maintain the robot active to give a sense of naturalness. In this context, the *Motivational model* presented in Chapter 6 is fundamental to generate a lively and natural behaviour based on the artificial internal state emulated in Mini.

- We believe that DMS has to include planned activities for the user. To satisfy this necessity, we designed an agenda (see Section 5.2.2.3) to store structured information about events that the user has to execute and reminders that the robot has to notify the user.

- The information stored in the robot's *Memory* contains personal data about each user of the robot. Thus, if the robot knows the user it is interacting with, we pretend that it adapts and customises its behaviour to the user to maintain engagement and improve the user perception of the robot.

- In the situations where the robot is motivated to play or the user wants the robot to execute an entertainment activity, the selection of the entertainment activity should be personalised to the user's features. For this reason, using the information stored in the user profile and the Preference Learning (PL) system, the robot has to maintain the behaviour personalisation by suggesting each user their preferred activities.

- The system must combine the autonomous execution of planned activities with adaptive and reactive behaviour to unexpected situations. We believe that this method improves the naturalness of the robot and allows the user to perceive what the robot is experiencing.

- The DMS should be able to change between behaviours dynamically. Moreover, the DMS must give the users the possibility to cancel an ongoing activity and start the activity they prefer. We propose that every time the user wants to cancel an ongoing robot's behaviour, they have to touch its right shoulder or say the robot's name. Then, the robot asks users if they want to continue with the previous activity or perform a new one. If the user responds to continue, the activity is resumed. Otherwise, the activity stops, and the robot asks the user which activity they want to execute next using the tablet menus or voice commands.

7.2. The operation of the Decision-making System

The decision process involves a sequence of steps that evaluate the robot's situation and the inputs received from other modules. Figure 7.2 shows a diagram flow of the decision-making loop proposed in this contribution.

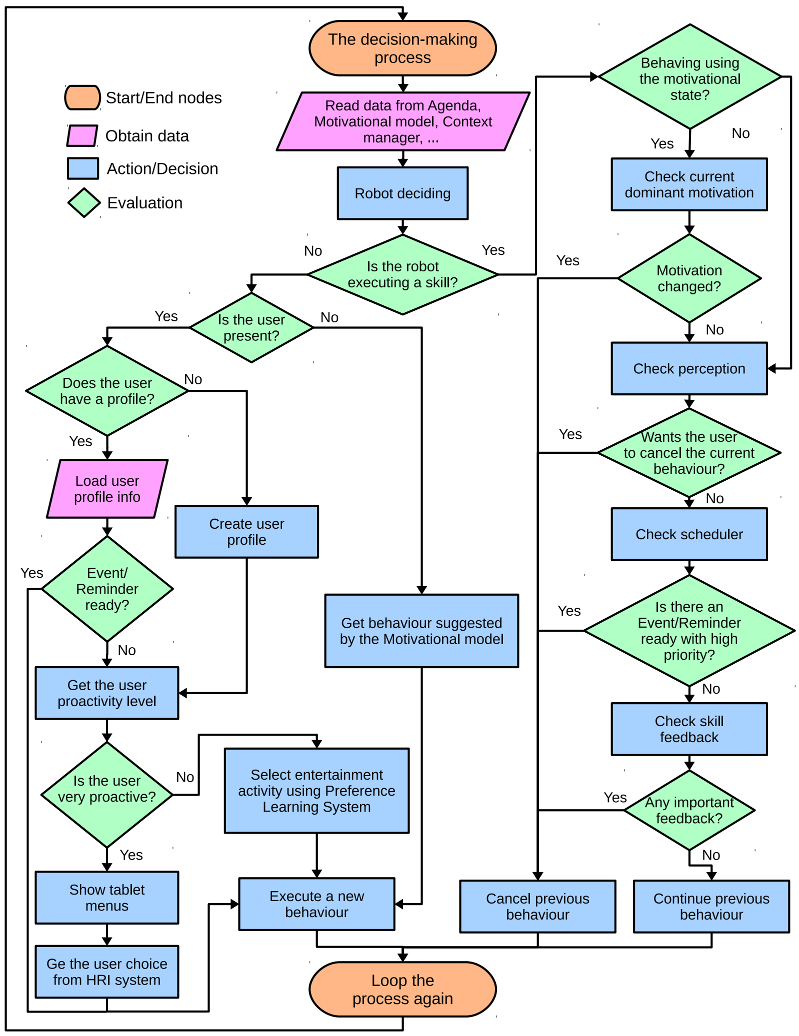


Figure 7.2: The decision-making loop proposed in this thesis.

The behaviour selection process carried out by the autonomous DMS presented in this thesis is as follows.

1. First, the decision module receives and evaluates the information from the external inputs such as the *Motivational model*, the *HRI manager*, the *User*

profiling manager, the *Scheduler manager*, and the *Context manager*. Using this information, the DMS starts making its decision if there is not an active behaviour.

2. Then, the system checks whether there is an active behaviour. If a behaviour is running, the system evaluates if there is a reason to stop it. Otherwise, the system evaluates if the robot perceives a user ready to interact with it.

3. If there is a user ready to interact with the robot, the robot checks, using the communication with the *Context manager* whether its profile is already created. If the profile is not created, the robot creates a new profile activating the *User profiling skill*.

4. If the user profile exists in the *Context manager*, the robot loads the user's information to personalise the interaction, including the user proactivity level.

5. If there is an event/reminder in the user's agenda that must be executed, the robot directly starts the related activity without the user intervention.

6. If there is no event/reminder near, the robot follows two possible alternatives: ceding the initiative to the user or making an autonomous selection. This decision is performed by assigning a probability to each alternative based on the user proactivity level and applying randomness. If the user proactivity is high, the robot is more likely to display menus on the tablet where the user can select the activity using the screen or voice commands. Otherwise, if the user proactivity is low, the robot is more likely to make an autonomous decision of the entertainment activity it will propose to the user. The autonomous decision is produced using the PL System presented in Chapter 9. The autonomous selection of the entertainment activity considers the user's preferences. It is carried out in three hierarchical steps (category, subcategory, and final activity) to improve exploration and balance between user-robot decisions using Boltzmann's equation described in Appendix B.

7. If there is not a user ready to interact with the robot, the robot behaves using the motivational state of the robot and the behaviour-based RL learning system. The robot decides which behaviour it will execute. For example, if the motivational state is sleep, the robot will go to sleep, and if the robot's motivational state is to play alone, the robot will dance.

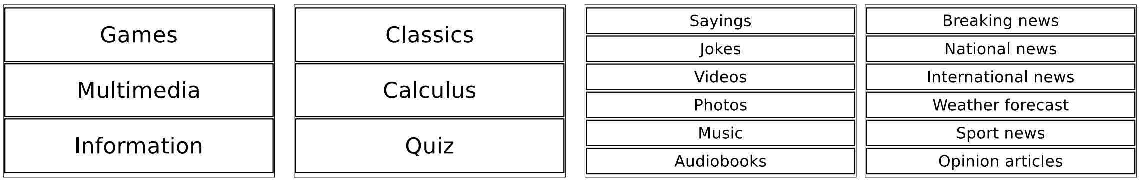
8. If any motivational state is active (the *Motivational model* fails to provide information to the *Central controller*), the robot will relax performing the waiting behaviour.

9. If the robot is executing a behaviour, the DMS evaluates the following inputs and respond to them when necessary:

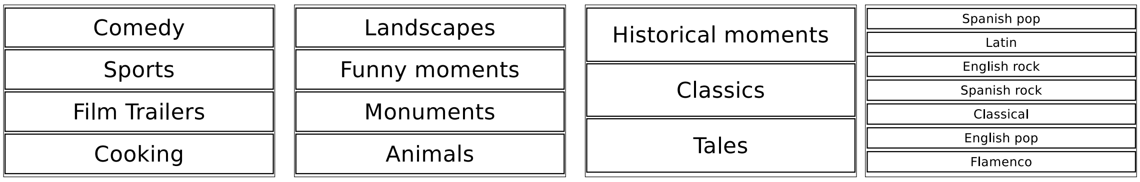
- If the robot is acting according to a dominant motivation and it changes, the DMS will cancel the current activity and start a new one in line with the new dominant motivation.
- If the user calls the robot's attention by touching his right shoulder or saying its name, the current activity is paused. Then, the robot asks the user whether to continue with the activity or perform a different one. If the user decides to cancel the activity, the *Central controller* stops the skill, and a new tablet menu appears. Using these menus, the users can select the activity they desire on the screen or by voice commands. Figure 7.3 shows the sequence of tablet menus displayed to select the desired activity. These menus organise the entertaining activities in categories and subcategories.
- If the user is not engaged skipping answering the questions of the robot, the *HRI manager* informs the *Central controller* about this issue, cancel the current activity. Then, the robot enters in waiting mode.
- If the *Scheduler manager* informs the DMS that a reminder will be notified to the user with high priority, the current behaviour is paused, the skill notifies the reminder, and then the robot resumes the previous activity.
- If the *Scheduler manager* informs that a new event has to be executed with high priority, the previous activity is cancelled, and the event is executed. If the event priority is low, the event is queued to be executed after finalising the current activity.
- It is essential to mention that when a skill is working and especially when it finishes, the feedback that it sends back to the DMS can indicate which action to do next. These kinds of feedback are limited to a predefined list of commands stored in configuration files. These files indicate the DMS what skill to execute, how to execute it, and its priority. An example of this case could be when the users watch photos using the tablet as the Tablet menus skill indicated the DMS that the following activity to run is the Photos skill.

10. Occasionally, during the standard operation of the system described in the previous steps. the robot inserts general questions to chat with the user and avoid fatigue after performing many consecutive activities.

This process repeatedly loops, allowing the robot to select the proper behaviour dynamically. Note that the behaviour selection process combines a motivational bio-inspired behaviour with adapting to the user's features to maintain engagement.



(a) Entertainment menu. (b) Games menu. (c) Multimedia menu. (d) Information menu.



(e) Videos menu. (f) Photos menu. (g) Audiobooks menu. (h) Music menu.

Figure 7.3: Tablet menus shown by the robot giving the users the possibility to select the option they prefer. The navigation through the menus can be done touching the screen or by voice commands.

7.3. Conceptual design

Figure 7.4 shows the DMS we propose in this thesis. It consists of two levels that exchange information to make appropriate decisions and control the robot's skills, that situate in another level outside our DMS. On the top of the DMS, the most critical module is the *Central controller*. It receives all the information from the different external subsystems, integrates it, and yields appropriate responses to control the different functionalities of the robot. Below the *Central controller*, the DMS has a lower level formed up by a group of five *Managers* called *Entertainment*, *Relax*, *Cognitive stimulation*, *Social*, and *General*. The *Central controller* activates a specific manager depending on the skill (activities) it decides the robot has to execute. Each *Manager* handles a different group of skills. Thus, for example, if the robot decides to play with the user, the *Entertainment manager* will be activated, whilst if the *Central controller* decides the robot has to sleep, it will activate the *Relax manager*. The inclusion of the *Manager* level in the architecture has a triple aim. First, the *Managers* alleviate the payload of the *Central controller* in managing such large number of instructions. Second, the *Manager* level abstracts specific functionalities depending on the kind of application that the robot is executing. This case is the integration of the PL module in the *Entertainment manager* for personalising the selection of the user's favourite entertainment activities. Finally, it simplifies control of simultaneous skills since our system allows to combine the execution of long-lasting activities such as games of multimedia displaying with short skills like reactions or communications to the user. In this situation, the *General manager* always controls the short skills and another *Manager* (*Entertainment*, *Relax*, *Social*, or *Cognitive stimulation*) controls the long-lasting skill.

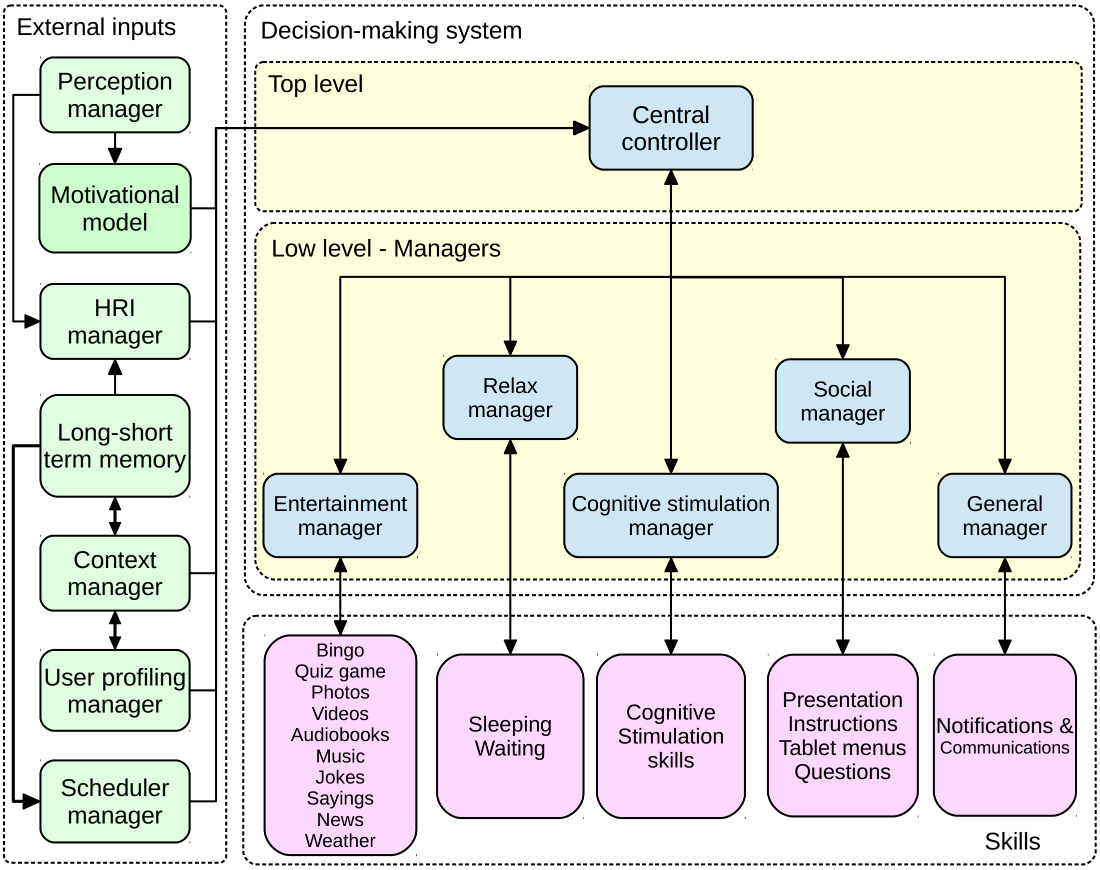


Figure 7.4: The DMS and its communication with the external managers. The DMS has three levels. In the top one, the *Motivational model* outputs the motivational state of the robot. At the medium level, the *Central controller* receives input information from other modules in the architecture to make appropriate decisions. The lower level cluster five *Managers* (entertainment, social, general, relax, and cognitive stimulation) control different skills depending on their functionality.

It is worthy of mentioning that although the DMS has two levels, their common purpose is to appropriately control the robot skills to deploy all the robot's functionalities. The *Skill* level situates outside the DMS as a layer that comprises all the robot's functionalities. Thus, as will be described later in this dissertation, the skills are not only controlled actuation commands but contain the internal logic of the function that the robot is performing, principally managing complex HRI scenarios. Next, we deepen into the description of the specific operation carried out by the *Central controller*, the *Managers*, and the *Skills*.

7.3.1. Central controller

The *Central controller* is a module inside of the DMS in charge of receiving information from multiple external modules of the robot (*Motivational model*, *HRI manager*, *Context manager*, *Scheduler manager*, and *User profiling manager*) for making a decision about how the robot has to behave. Once the *Central controller* makes a decision, it activates the corresponding *Manager* for executing the functionality decided. The *Central controller* and the *Managers* exchange information to maintain synchronous bidirectional communication and be aware of possible errors in both modules. In our system, the *Central controller* can simultaneously handle two *Managers* at a once to execute two simultaneous skills. As we mentioned before in this chapter, one of these *Managers* must be the *General manager* controlling a short-lasting skill while the other skill is controlled by one of the other *Managers* in the DMS (*Entertainment*, *Relax*, *Social*, or *Cognitive stimulation*).

7.3.2. Managers

The function of a *Manager* is to receive the commands from the *Central controller*, process them, activate the demanded skill, and control its execution. Most of the time, the *Central controller* decides the skill and the specific function (goal) that the robot has to execute. Next, we define the specific functionality of each Manager in the DMS.

- **Entertainment manager:** The functionality of the *Entertainment manager* is to control those skills used for entertaining the user. Unlike the other *Managers*, this one can make partial decisions about which activity the robot will execute using the Preference Learning (PL) system designed to learn the user preferences (see Chapter 9). The inclusion of the PL in the *Entertainment manager* instead of in the *Central controller* allows alleviating the processing charge of the *Central controller*.

- **Social manager:** The *Social manager* controls those functionalities that imply verbal and non-verbal communication with the user, like questions or chatting.

- **Relax manager:** The *Relax manager* controls the sleeping and waiting skills. The sleeping skill simulates that the robot is sleeping, while the waiting skill simulates that the robot is resting but awake.

- **Cognitive stimulation manager:** The *Cognitive stimulation manager* is a particular case since it only manages the cognitive stimulation skills (personalised treatments, planned sessions, and individual exercises). These functions handle the therapies that patients with cognitive impairment must follow, being part of another thesis [302] carried out previously in our research group.

- **General manager:** This manager controls general functionalities of the DMS like asking the user about continuing with an action, notifying reminders to the user or informing the about the next activity that it will execute. The user of this *Manager* is necessary to combine the execution of two skills simultaneously.

As mentioned before, the architecture allows that more than one *Manager* activates simultaneously, each one controlling one skill from its group. However, the system does not allow that the same *Manager* controls two skills simultaneously since each skill deploys a full functionality that captures several robot interfaces simultaneously. Considering this fact, the *General manager* can be simultaneously activated with one of the other *Managers*. This activation is possible because the *General manager* handles short functionalities that only interrupts other skills for a few seconds. The following section addresses the skill level, a subsystem out of the DMS but closely related to it.

7.3.3. Skills

The robot's skills allow it to deploy different functionalities towards completing the task it has to attain in every moment. As mentioned in Section 5.2.3, the robot has two different types of skills: continuous and conditional. Continuous skills are those that are always active and ready to work. Two examples of continuous skills are the TTS and ASR, two functionalities that are always ready for maintaining verbal communication with the user. Otherwise, conditional skills activate under the control of the DMS in particular circumstances, being turned off the rest of the time. Examples of this type of skill are the weather forecast, the cognitive stimulation exercises, or the quiz game. Note that while it is infrequent to have more than one conditional skill active simultaneously, all the continuous skills work simultaneously. It is essential to mention that the contribution of designing continuous and conditional skills was carried out in a previous thesis of the group [19]. Nevertheless, in this work, we have extended the functionality of conditional skills by allowing them to be paused and resumed while sending periodic feedback to the DMS. Bidirectional communication allows a synchronised exchange of information that strengthens the system operation, including skill control.

In this thesis, we concentrate on the control of the conditional skills since they are the ones managed by the DMS. The modularity and flexibility of the skill network allow developers to design new skills without altering the operation of the rest of the system. Thus, we can extend the robot's functionalities according to our necessities. As Figure 7.5 shows, each conditional skill works as a state machine whose state is indirectly controlled by the *Central controller* throughout its action on the correspondent manager. A conditional skill can be only in one state at each step (started, completed, paused, cancelled, resumed, or stopped), working differently depending on each situation. It is essential to mention that skills can perform different functionalities depending on the goal message that the manager sends to them. For example, the music skill can play different music styles, or the News skill can inform different information. It is also worth mentioning that we call activities or applications to the different variations that skill can take in this work.

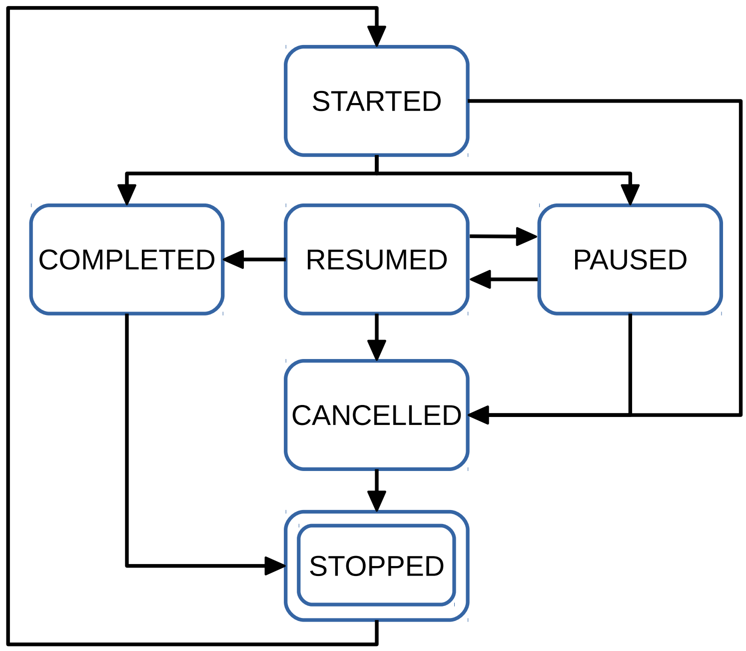


Figure 7.5: State machine describing the operation of a conditional skill.

Next, we describe the skill's operation depending on its active state.

- **Started:** The skill enters in this state when the DMS activates it. Then, the skill waits to receive the goal to execute and performs its main functionality.
- **Paused:** The DMS can pause the skill if it has to perform another activity. This state normally activates in two cases: if the robot has to communicate something to the user, in which case the skill will later continue from its previous saved point after the communication, or when the user calls the robot to stop the skill, in which case the robot asks the user if he wants to continue with the skill or execute another operation.
- **Resumed:** This state activates only after a pause, meaning that the robot continues executing the active skill. Like in the started state, the robot can fulfil its goal.
- **Completed:** The skill enters in this state when it finishes the goal that it received from the DMS. This state is the only one whose activation does not depend on the DMS since the completion of the tasks is handled autonomously by the own skill.
- **Cancelled:** The skill enters in this state when the *Central controller* aborts a goal. This fact can occur if the user voluntary stops the skill or if the DMS decides its time to perform another activity.
- **Stopped:** Conditional skills are by default in this state, meaning that they are not active.

Each skill belongs to one of the *Managers* listed above, depending on the functionality it performs. Next, we enumerate the skills controlled by each *Manager* in the DMS architecture.

- **Skills controlled by the *Entertainment manager*:** The group of skills comprises the Bingo, Quiz game, Photos, Videos, Audiobooks, Music, Jokes, Sayings, News, Dancing, and Weather skills, all of them oriented to entertain the user.
- **Skills controlled by the *Social manager*:** The *Social manager* includes the Presentation skill, the Instructions skill, and the skill that controls the tablet menus and general questions to the user.
- **Skills controlled by the *Relax manager*:** In this group, it is possible to find the Sleeping skill and the Waiting skill.
- **Skills controlled by the *Cognitive stimulation manager*:** This group of skills includes personalised treatments for each user, daily sessions for working in different areas (e.g. memory or perception), and individual exercises.
- **Skills controlled by the *General manager*:** These skills are those used for communicating general information to the user (Notifications, Alerts and Communications skills).

7.4. Implementation

The following sections explain the implementation of the DMS into the robot software architecture described in Section 5.2. The description focuses on how the *Central controller* internally communicates with the *Managers* and receives external information from the external modules in the robot, and how the *Managers* communicate with the *Skills*. As mentioned previously in this thesis, the implementation lies on Robot Operating System (ROS), a set of software libraries and tools that facilitate the creation of robotic applications. Figure 7.6 shows the Robot Operating System (ROS) system in terms of nodes (modules), messages with topics, and services inside the DMS and between the *Manager* level and the *Skills*. The messages and services used in the communications are custom and designed for our software architecture.

7.4.1. Communication between the Central controller and the Managers

On the one hand, the *Central controller* sends a custom *ManagerControl* ROS message containing information about how the manager has to behave. The DMS sends this message every time it has to update the status of a *Manager*. The *ManagerControl* message contains the following attributes:

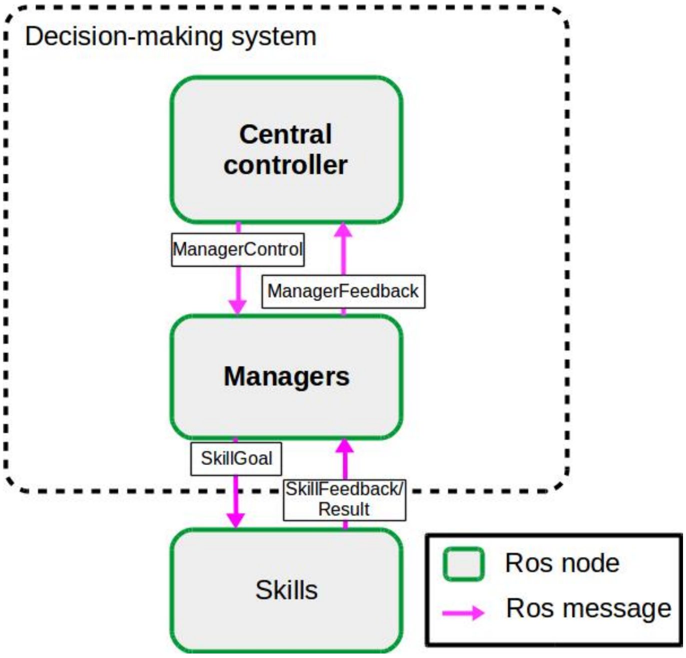


Figure 7.6: Internal communications in the DMS using ROS.

- **Manager:** Name of the manager in charge of controlling the execution of the skill.
- **Action:** Name of the skill that the manager has to control.
- **Goal:** Name specifying the functionality that the skill has to execute.
- **Skill status:** Command given to the skill to control its internal state machine.
- **Control parameters:** List of additional key-value parameters necessary for the execution of the functionality.

Table 7.1 shows an example of a message sent by the *Central controller* to handle a *Manager* and start a *Skill*.

Parameter	Value
Manager	Entertainment
Action	Music
Goal	Latin
Skill status	Start
Control parameters	-

Table 7.1: *ManagerControl* message example.

On the other hand, every time the manager or the skill updates its state. It sends a custom *ManagerFeedback* ROS message containing information about the *Manager* and the skill. This message contains the following attributes:

- **Manager:** Name of the manager controlling the execution of the skill.
- **Action:** Name of the manager's skill to control.
- **Goal:** Name the goal (specific functionality) that the skill has to execute.
- **Manager status:** Status of the manager while controlling the current active skill.
- **Skill status:** Status of the skill that the manager is controlling. The status can be started, completed, paused, resumed, cancelled, or stopped, as Figure 7.5 shows.
- **Terminal status:** Final result of the skill once it has finished.
- **Engagement:** Flag indicating if, while performing the functionality, the user is engaged with the robot or not. This variable measures if the user responds to the robot's questions.
- **Response parameters:** List of additional key-value parameters that the skill sends to the *Central controller* due to the interaction of the user and the environment.

Table 7.2 shows an example of the ManagerFeedback message used to inform the *Central controller* about the status of a *Manager* and the *Skill* it is controlling.

Parameter	Value
Manager	Entertainment
Action	Music
Goal	Latin
Manager status	Running
Skill status	Started
Terminal status	None
Engagement	True
Response parameters	-

Table 7.2: *ManagerFeedback* message example.

7.4.2. Communication between a Manager and a Skill

The communication between an active skill and its correspondent *Manager* is synchronous and bidirectional. This fact means that every time the manager sends an instruction to the skill, the skill has to immediately respond, confirming the reception of the command and information about its state. Besides, the skill sends periodic feedback about its operation, the interaction with the user (it exists), its state, and the result of the goal indicated by the *Central controller*.

The communication between each skill and the manager works with Actionlib[3], a ROS package that provides a robust interface for defining customised actions with

[3]http://wiki.ros.org/actionlib

predefined goals using a client (manager)-server(skill) communication. The Actionlib service contains three fields: goal, result and feedback. The goal is the specific task (activity) that the skill has to fulfil. It can contain many parameters, but we opted to standardise them in this thesis using the following ones for the *SkillGoal*.

- **Skill command:** String that defines the goal that the skill has to execute.
- **Max time:** Maximum duration of the skill. If set to 0, the skill will finish when the skill reaches a solution, or the DMS cancels it.
- **Number of plays:** Number of times that the goal loops. Useful for skills like playing music or games.
- **User proactivity:** The level of user proactivity from 1 to 5. A value of 1 defines a user that does not take the initiative, whilst a value of 5 indicates a very proactive user. In our DMS, this parameter is noteworthy since it balances how the robot's addresses the autonomous selection of entertainment activities. On the one hand, if the user's proactivity level is high, the robot will, with high probability, cede the initiative to the users to select their preferred entertainment activity. On the other hand, if the user proactivity is low, the robot will make autonomous proposals more likely than ceding the user's initiative. We balance between exploring unusual selected actions and the user's favourite alternatives using this method. If the user proactivity is moderate, the chances of making an autonomous decision or ceding the initiative to the user are equal.

Table 7.3 shows an example of a goal sent by a *Manager* to control a *Skill* that is going to be executed.

Parameter	Value
Skill command	Start
Max time	10
Number of plays	1
User proactivity	3

Table 7.3: Message from a *Manager* to control a *Skill*.

Each skill provides the manager with periodic feedback every time a new command confirms its reception. As with the goal, all the architecture skills have the same feedback fields. The skill updates the percentage completed and if the user is not engaged with the robot. The *SkillFeedback* message contains the following parameters.

- **Skill result:** Result of the skill. It is sent once after the execution of the skill. It would have a value of 0 if the result is successful, -1 if there was an error during the execution, or 1 if the goal failed due to a user or DMS cancellation.

- **Skill status:** Current status of the skill. This variable contains one of the states set in Figure 7.5.
- **Percentage completed:** Percentage completed of the skill from 0 to 100. The skill internally manages the value that sends the feedback every time it changes.
- **Engagement:** Flag indicating if the user is engaged or not with the robot. By default, the flag value is True, but when the user does not answer a question or ignores the robot, its value changes to False.
- **Response parameters:** Additional configuration parameters that conditional skills send to the DMS informing, for example, what skill the robot has to do next. This feedback field is formed by key-value pairs containing many different commands. These orders are called DMS responses and are predefined in a configuration file.

Table 7.4 shows an example of the message returned to a *Manager* by one *Skill* after finishing.

Parameter	Value
Skill result	0
Skill status	Completed
Percentage completed	100
Engagement	True
Response parameters	-

Table 7.4: Feedback from a *Skill* to a *Manager*.

Finally, every time the skill finishes its execution, it has to send the manager information about its result, using the message *SkillResult*. Like in the previous cases (goal and feedback), the standard result message contains a single parameter called skill result whose possible values are -1 if the skill returned an error, 0 if the skill finished successfully and attained its goal and 1 if the skill finished successfully, but the goal was not attained.

7.4.3. Communications with the external modules with influence on decision-making

The *Central controller* receives many external inputs from other subsystems of the robot to support the decision-making process. Then, the following sections describe the kind of information that each of these subsystems in the architecture sends to the *Central controller* module. Figure 7.7 shows the communications using ROS messages and services between the external modules of our robot architecture and the *Central controller* of the DMS.

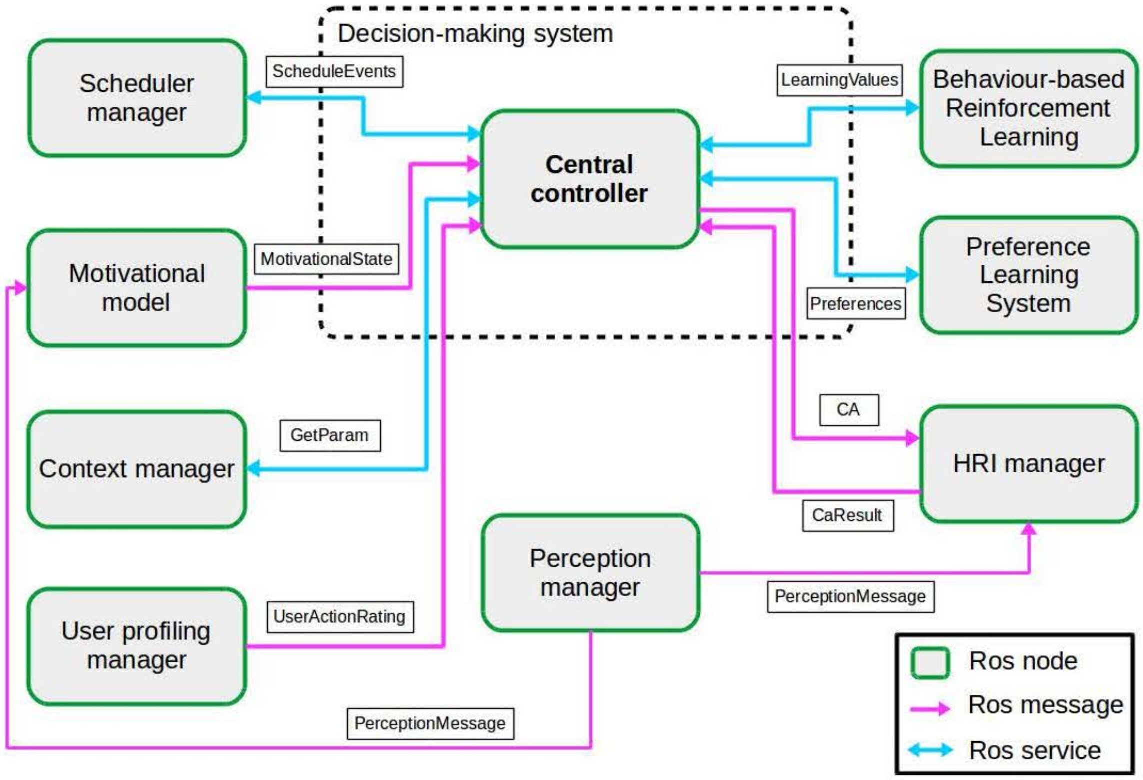


Figure 7.7: Communications of the Central controller of the DMS with external modules of the robot architecture.

7.4.3.1 Communication between the Central controller and the Motivational model

The communication between the *Motivational model* and the *Central controller* is carried out using our *MotivationalState* custom ROS message that sends the DMS the robot's motivational state. This message contains the following attributes:

- **Dominant motivation:** The name of the motivation with the highest level of intensity among all active motivations.
- **Behaviour:** Recommended behaviour to execute according to the internal state of the robot, which is represented through the dominant motivation.
- **Motivational intensities:** Array with the names and intensities of all the motivations included into the *Motivational model*.

Table 7.5 shows an example of the message sent from the *Motivational model* to the *Central controller* containing the robot's motivational state.

Parameter	Value
Dominant motivation	Relax
Behaviour	Sleep
Motivational intensities	Relax: 78, Play: 24, Socialise: 12...

Table 7.5: Example of the message exchanged between the *Central controller* and the *Motivational model* containing the robot's motivational state.

7.4.3.2 Communication between the Central controller and the HRI manager

The communication between the *Central controller* and the *HRI manager* works under the supervision of the *Perception manager* in two steps. On the one hand, the *Perception manager* sends unidirectional periodic information to the *HRI manager* informing about changes in the environment that influence the decision-making process.

The function of each of the parameters of the *PerceptionMessage* is as follows:

- **Header:** The header contains basic information about the message, like the ROS node that is sending the message and the time it was published.
- **Abstraction level:** The abstraction level indicates the level of the *Perception manager* that is sending the message. This variable is typically set to 0 as the messages come from the Level 0 in the *Perception manager*.
- **Values:** List of key-value pairs containing meaningful information about the perception like the status of the stimuli and its intensity.

Table 7.6 shows a communication example between the *Perception manager* and the *HRI manager*.

Parameter	Value
Header	26/01/2022 15:33:21
Abstraction level	0
Values	Item: Touch, Intensity: 76, State: Right shoulder

Table 7.6: Example of the message exchanged between the *Perception manager* and the *HRI manager*.

At the same time, the DMS requests the *HRI manager* to capture and be aware of changes in the touch sensors and in the voice recognition system to know when the user is calling the attention of the robot. In our system, the user can call the robot's attention using two different mechanisms: touching its right shoulder and calling it by its name. In case the *HRI manager* informs the DMS that any of these events have happened, the DMS will give the initiative of the interaction to the users, letting them select the activity that they prefer. Next, we describe the field of the *CaResult* message provided by the *HRI manager* to the DMS providing information about the interaction of the robot with the environment.

- **Header:** This field contains basic information about the message, like the ROS node is publishing the message or the timestamp.
- **Emitter:** Indicates which ROS node requested the information to the *HRI manager*.
- **Result:** Result of the petition. This field indicates if the response status was success or fail.
- **Values:** This field contains the response of the *HRI manager* about the petition like, for example, where the user touched the robot.

Table 7.7 shows an example of a message received by the *Central controller* from the *HRI manager*.

Parameter	Value
Header	08/09/2021 11:24:12
Emitter	Decision-making
Result	Success
Values	Answer id: touch, Answer time: 12.23, Answer value: Right shoulder...

Table 7.7: Example of the message exchanged between the *Central controller* and the *HRI manager*.

7.4.3.3 Communication between the Central controller and the User profiling manager

The *User profiling manager* communicates with the DMS providing information about how each specific user is interacting with the robot, as Section 5.2.2.1. In the *User profiling manager*, a unique user profile is stored for each user that interacts with the robot. Then, the *User profiling manager* gathers interaction metrics after the execution of each activity to update the user profile. The *User profiling manager* builds a set of tables containing information about how each user interacts while executing each particular activity.

The following list shows the *User profiling manager's* interaction metrics for each particular activity that the robot can execute. The *Manager* stores the metrics for the last time an activity was executed and the average metrics of all the activity was executed. Besides, for each activity, the system stores the average values of all the users. These metrics are provided using the *UserActionRating* ROS service under petition of the *Central controller*.

- **Manager:** Name of the manager.
- **Action:** Name of the skill.
- **Goal:** Name of the goal.
- **Start:** Starting date-time in string format.
- **End:** Finishing date-time in string format.
- **Duration:** Duration of the activity.
- **Terminal status:** Result of the activity (succeeded, failed, or error).
- **Engagement:** Flag indicating if the user was engaged during the interaction with the user. This variable measures if the user is responding to the robot's questions during the interaction.
- **Correct answers:** Number of correct answers.
- **Wrong answers:** Number of incorrect answers.
- **No responses:** Number of no responses.
- **Total answers:** Number of total answers summing up correct and wrong answers and no responses.
- **Answer time:** Average time the user takes to answer a question.
- **Interactions:** Number of interactions during the activity.
- **User selection:** Flag indicates whether the user selected the activity (true) or by the robot (false).

Table 7.8 shows an example of the interaction metrics sent from the *User profiling manager* to the *Central controller* to adapt the interaction with the user.

Parameter	Value
Manager	Entertainment
Action	Photos
Goal	Animals
Start	01/01/2022 08:56:18
End	01/01/2022 08:59:33
Duration	195
Terminal status	succeeded
Engagement	True
Correct answers	0
Wrong answers	0
No responses	0
Total answers	2
Answer time	13.24
Interactions	2
User selection	False

Table 7.8: Example of the interaction metrics sent by the *User profiling manager* to the *Central controller*.

7.4.3.4 Communication between the Central controller and the Context manager

As mentioned in Section 5.2.2.2, the *Context manager* updates personal information from each user for adapting the interaction. In the system, the *Context manager* and the DMS communicate throughout the action of the robot *Memory*. The *Memory* works as a kind of board that stores information that all the system ROS nodes can read from using the GetParam ROS service. In this case, the DMS reads the following information from the robot's *Memory*.

- **User proactivity:** The user proactivity indicates how active the user is while interacting with the robot. It is used to promote the interaction if the value is too low or let the user take the initiative if it is high. The value ranges from 0 – 5.
- **Personal information:** The system reads from the parameter server personal information like the user's name, location, or age to produce a personalised interaction experience.
- **Last executed activities:** The last executed activities allows the system to avoid the repetition of the ones that have been recently executed, promoting the exploration of new activities.

Table 7.9 shows an example of the parameters and their values gathered by the *Central controller* from the *Context manager* for adapting the interaction.

Parameter	Value
User proactivity	4
Personal information	Name: Paul, Age: 23, ...
Last executed activities	Videos: Cooking, Games: Quiz game

Table 7.9: Example of parameters gathered from the *Context manager* by the *Central controller*.

7.4.3.5 Communication between the Central controller and the Scheduler manager

The *Scheduler manager* reads and organises the events and reminders stored in the *Context manager* that each specific user has to execute, as described in Section 5.2.2.3. The *Scheduler manager* and the DMS exchange information under the petition of the latter to produce an appropriate notification of the reminders and the events. The message contains two sections, the request and the response. The DMS requests the *Scheduler manager* which events and reminders should be executed using the ScheduleEvents ROS service. The *Scheduler manager* returns if there is a new event-ready, the arguments (using key-value pairs) of the next event or reminder that the DMS has to perform.

Each event/reminder has a set of parameters containing the information for the DMS to execute it.

- **Item id:** Unique id for each event/reminder.
- **Status:** Current status of the event (confirmed, tentative, or cancelled.
- **Type:** Type of the event (reminder or activity).
- **Manager:** Manager where the activity is classified. In case it is a reminder, the manager is always the general one.
- **Action:** Name of the skill to execute
- **Goal:** Name of the goal to execute
- **Start:** Date-time of starting.
- **End:** Date-time of ending.
- **Created:** Date-time when the event was created.
- **Duration:** Duration of the event.
- **Priority:** 3 (high), 2(medium), and 1(low).
- **Recurring:** Flag indicating if the event repeats over time or just once.
- **Periodicity:** Period of repetition. Ignored if the previous parameter is set to false.

- **Description:** Brief description of the event.
- **Text:** Text to communicate in the reminder.

Table 7.10 shows an example of a reminder sent from the *Scheduler manager* to the *Central controller* for its execution.

Parameter	Value
Item id	002
Status	Confirmed
Type	Reminder
Manager	General
Action	Notifications
Goal	Reminder
Start	10/12/2021 14:00:00
End	-
Created	08/12/2021 18:22:03
Duration	-
Priority	1
Recurring	False
Periodicity	-
Description	Appointment with the doctor
Text	I would like to remind you an appointment with your doctor

Table 7.10: Example of a message containing a reminder.

7.4.3.6 Communication between the Central controller and the Preference Learning system

The DMS adapts its action selection depending on the user it is interacting with. The robot uses their preferences towards the entertaining activities of the robot to decide the best alternative. The PL framework works in two stages. The first stage consists of estimating the user preferences using a machine learning algorithm (review Section 4.3.2) that generates the prediction using the user's characteristics and comparing them with the preferences of similar users. A database stores their information. The second stage consists in adapting the initial estimation with the HRI. Every time the robot wants to suggest to the user an entertainment activity, the DMS requests the Preference Learning (PL) module the user preferences using the custom *Preferences* service. This service contains the following information.

- **Activities:** Structured array with the names of the activities organised in ranking format.

- **Values:** Score assigned to each activity contained in the previous array. These values are sorted using the same order as the Activities array.

Table 7.11 shows an example of the previous message during the communication of both modules.

Parameter	Value
Activities	Quiz game, Bingo, Calculus game, ...
Values	3.12, 2.78, 1.12...

Table 7.11: Example of the message exchanged between the Central controller and the PL System.

7.4.3.7 Communication between the Central controller and the Behaviour-based Reinforcement Learning system

The *Motivational model* provides the DMS with information about the emulated needs, intentions, and desires of the robot. Initially, the robot does not know how to behave to maintain an optimal internal state since the only information it has is the effect of each behaviour on the hormones, neurotransmitters, and biological processes. Despite in Table 6.7 each motivational state ties to a specific behaviour, the robot initially lacks this information. For this reason, we designed a Behaviour-based RL system that works in parallel with the *Motivational model* and allows the robot to map its motivational state to behaviours (see Chapter 10). This information is provided every time the robot behaves using its motivational state using the *LearningValues* ROS service. This service contains the following attributes.

- **Behaviours:** List of behaviours of the robot.
- **Q-values:** Associated Q-values (scores) to each behaviour listed in the Behaviours array.
- **Counts:** Array containing the number of times each behaviour has been executed.
- **Robot state:** String defining the current internal and external state of the robot.

Table 7.12 shows an example of the information sent by the Behaviour-based RL System to the *Central controller* to decide the most appropriate behaviour.

Parameter	Value
Behaviours	Play, Sleep, Talk, Dance, ...
Q-values	10.34, 7.33, 12.3, -1.23, ...
Counts	21, 24, 25, 31, ...
Robot state	Socialise, user present, music player off

Table 7.12: Example of the message exchanged between the Central controller and the Behaviour-based RL system.

7.5. Conclusion

This chapter has presented our DMS, the cornerstone of this dissertation. As we have presented, this module is essential for reaching an autonomous and natural behaviour during long-term HRIs since it produces the most appropriate behaviour after evaluating the current situation that the robot is experiencing.

Following, we present the results derived from the realisation of this thesis, explaining the methodology, setup, and evaluation we followed in carrying out each of the experiments that served us to validate the performance of our system.

Part III

Experiments and results

This part describes the experiments conducted to test and assess the operation of the DMS and its functionalities developed in this thesis. It is divided into the following chapters.

- Chapter 8 presents the research work carried out during the stay at the University of Hertfordshire in Hatfield, United Kingdom related to robot adaptation. The work consisted of designing an adaptive system for the socially assistive robot MiRo to dynamically shape its affective state according to the quality and procedures of the interaction with people. To assess the system, we defined various user and robot behavioural profiles that were compared against each other to notice variations in the robot's behaviour. The experiments were carried out both in simulation and real HRIs.

- Chapter 9 describes and presents the results of the PL system for personalising the robot's entertainment activities selection. The learning system intends to serve as a recommendation system that balances the autonomous selection of the users' favourite activities and the execution of unexplored activities to build up the users' preferences. The validation of the system consisted of two steps. In the first place, we developed a module that predicts the user preferences based on the user's features by comparing them with similar users stored in a dataset.

- Chapter 10 describes a machine learning system based on RL techniques to map the robot's motivations to actions considering the effects that each action produces on the robot's internal state. In an initial approach, we used Q-learning to allow the robot to learn while interacting with external objects and agents in the environment. Then, as Q-learning requires long-lasting training periods for large state-action spaces and is vulnerable to the duration of the actions, we moved to design a Dyna-Q+ approach based on a model of the environment that increased the learning speed and stability.

- Chapter 11 contains the experiments related to testing the operation of the motivational model presented in Chapter 6. Our motivational model aims to extend our vision to new approaches that deepen into the neuroscientific basis of human behaviour. Thus, in Section 11.1 we show the evolution of essentially biological processes primarily involved in motivation and affect as the origin of human behaviour and expressiveness. Our motivational model's improvements pretend to provide the robot's more varied, adaptable, and expressive behaviour allowing Mini to adapt to unexpected situations. Besides, the model shapes the robot behaviour for full-time operation and long-lasting interactions.

As the next step in modelling of biological functions to enhance the robot's behaviour and HRI, in Section 11.2 we present an application of the motivational model mentioned above that allows Mini to develop social bonds with the users it interacts. As an example of social adaptation, Mini perceives how good the user behaves and responds accordingly. If the user behaves kindly, Mini will execute social behaviours more often. Otherwise, the robot will avoid interacting with the user by performing escaping behaviours.

- Finally, Chapter 12 describes how we express affective states in Mini using the emotion and mood generated in the *Motivational model*. In this functionality, it is essential the role of the *Expression manager*, a module that receives the robot's affective state and modulates the robot's expressions. First, the chapter describes the system we have designed to control Mini's actuators using biologically inspired functions. Thus, using the neuroendocrine motivational model and the modelling of the autonomic functions previously presented in Chapter 6, the robot modulates its heart rate, bling rate, breathing rate, pupil size, and locomotor activity. Then, we validated the effect of showing emotions and mood on the robot's expressiveness by conducting a between-subject study to explore whether robot users perceive the emotion and mood expressed by Mini.

Next, in each of the chapters, we describe the methodology, experiment setup, evaluation, validation and present the results of the functionalities mentioned above devised to support the decision-making system of the robot. During the tests, more than 600 people participated by completing any of the surveys we used for the evaluation, and more than 100 people personally interacted with the robot in real HRI.

CHAPTER 8

Evaluating the robot's affective adaptation

This chapter describes the adaptive mechanism developed during the research stay in the Embodied Emotion, Cognition, and (Inter-)Action (EECAIA) Lab at the University of Hertfordshire. This work is an important source of inspiration for developing the mechanisms included in our DMS. This work aimed at studying the robot's affective behaviour regulation depending on the user it faced. We integrated our model into the social robot MiRo. The model simulates the robot's internal state with homeostatically controlled processes. The deficits of these processes urge the robot to reduce them executing certain behaviours to maintain an optimal internal state. Besides, the robot's affective state based on emotional responses modulates the behaviour expressiveness to show how it feels. We evaluated the model in HRI scenarios both in simulation and reality, comparing the robot behaviour when including and not including our proposed adaptive mechanism.

Next, we present the social robot MiRo, methodology, experiment, evaluation, and results of the work, ending with a discussion about the influence of this work on the rest of the manuscript.

8.1. The social robot MiRo

Unlike the rest of the results presented in this thesis, the experiments described in this section were carried out using the social robot MiRo [240], a commercial research platform for HRI. MiRo's appearance looks like a small rabbit that can wander in the environment seeking new stimuli that arouse it, as Figure 8.1 shows. In this work, we used its differential navigation and visual perception system based on two stereo cameras to navigate the environment, searching for users to interact with and objects to play with. Besides, we used the tactile sensors spread around its upper case to detect tactile contact as hits and strokes. The light sensors placed on its head helped MiRo perceive the room's illumination and whether it was night or day.

To express affective states, we use the actuators of the robot. On the one hand, we control the amplitude and speed of the motors that control the movement of the robot's head, neck, ears, and tail. On the other hand, we use the LEDs place on its shell to link each emotion with colour and let the user know how MiRo feels. Finally, we use its speaker to play non-verbal sounds simulating this kind of animal's sounds.

Figure 8.1: Front view of the social robot MiRo.

8.2. Robot architecture

Figure 8.2 shows the affective architecture we developed at the EECAIA for endowing MiRo with adaptive behaviour and affective expressiveness.

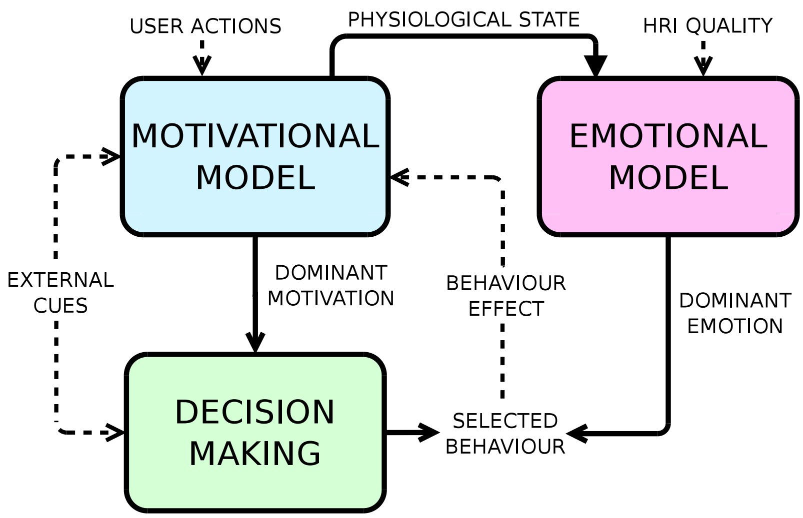


Figure 8.2: The proposed adaptive affective architecture integrated into the social robot MiRo.

The architecture has three main modules:

- **Motivational model:** This module simulates biological processes in the robot that are used to produce coherent decision-making. These processes vary with time, representing the robot's needs of sleeping, socialising, and entertainment and influenced by the user actions and external cues that define stimuli with the influence of the robot. Such deficits translate into specific motivations that urge behaviour.

- **Emotional model:** This module generates the robot's affective state depending on the interaction quality with the user and the robot's physiological state, affecting how the robot expresses its behaviour and current state.
- **Decision-making system:** This module receives the dominant motivation from the Motivational model and, after evaluating the robot's situations and external cues, decides the following behaviour. The robot's actions primarily focus on reducing the robot's deficits for maintaining the best possible internal state.

8.3. Motivational model

In the Motivational model proposed for this work, we decided to include three homeostatically controlled processes, each controlling a robot biological function:

- *Energy*: Represents the energy of the robot to execute actions. The *Energy* reduces when the locomotor activity of the robot increases.
- *Social*: Represents the robot's socialisation with people. Its value increases with time in the absence of socialisation. This variable derives from social needs.
- *Entertained*: This process represents the robot's entertainment. It reduces its value in the absence of entertainment activities. This value derives from the entertainment deficit.

The value of the homeostatic processes (h_i) ranges from 0 to 100 units, with an ideal value (iv_i) that represents where the state of the process is optimal, and its deficit satisfied. The value of the homeostatically controlled processes evolves following Equation 8.1. Each time step (0.2 seconds), the homeostatically controlled processes value h_i decrease by a decay rate δ_i. *Energy* has a decay rate of 0.2 units per time step, *Entertained* 0.5 units per time step, and *Social* 0.3 units per time step. The decay rate δ_i of each homeostatic process represents how fast a deficit appears in the agent. Figure 8.3 shows how the homeostatic processes emulated in MiRo decay with time.

$$h_i(t + 1) = h_i(t) - \delta_i \tag{8.1}$$

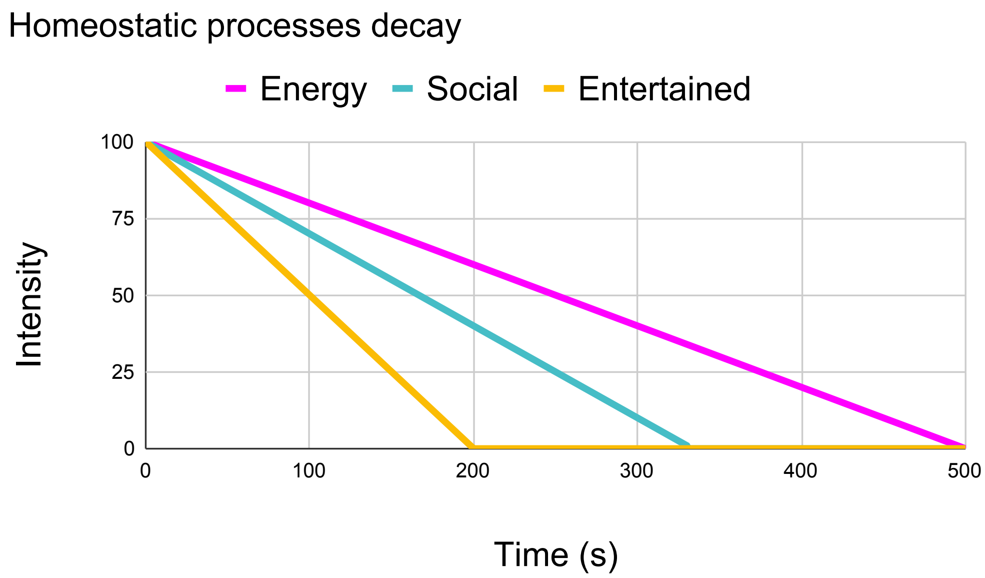


Figure 8.3: Decay with time of the homeostatic processes included in the social robot MiRo. Since we promote social behaviour, the Entertained and Social processes decay faster, so their deficits rise faster eliciting social behaviours more easily.

Table 8.1 shows the parameters that define the evolution of the homeostatically controlled processes in MiRo.

Process	**Ideal value $\mathbf{iv_i}$**	**Decay rate δ_i**
Energy	100	0.2
Social	100	0.3
Entertained	100	0.5

Table 8.1: Homeostatically controlled processes in MiRo and their decay rate δ_i.

When the value h_i of a homeostatically controlled process deviates from its ideal value (iv_i), a deficit (need) d_i emerges in the agent worsening its internal state as Equation 8.2 shows. The agent must execute the appropriate behaviour to reduce such deficits in each situation.

$$d_i(t) = |iv_i - h_i(t)| \tag{8.2}$$

The evolution of the robot's needs and behaviour depends on the stimuli it perceives and how it uses environmental resources. Thus, we define external stimuli like environmental changes that influence the robot's internal state. As explained previously, a significant deficit in the robot will drive it to perform actions to reduce it. However, there are some cases where stimuli can motivate an agent to behave differently even if the deficit is moderate. In our model, the external stimuli that influence the robot are the user's presence, level of illumination and coloured balls used for playing.

Each external stimulus has an associated perception signal ranging from 0 to 100 units representing the intensity of such stimulus. In this experiment, the robot's perception system provides the intensity with which each stimulus is perceived. Additionally to the influence of stimuli on the robot's deficits, stimuli are important regulators of motivational states, affecting the robot's decision-making and urging behaviour.

Motivational states depend on the deficit of a homeostatic variable to represent the robot's urges to reduce such deficits (d_i). Motivations are defined by an intensity level (m_i) calculated using a subtle variation of Equation 6.8. In this approach, the intensity of each stimulus si_k is regulated by an individual weight β_k and, at the same time, deficits (d_i) modulate how we perceive those stimuli, like Equation 8.3 shows. According to Equation 8.3, more than one stimuli (K) can influence a motivation, each with its corresponding modulating weight β_k and intensity si_k.

$$m_i(t) = d_i(t) + d_i(t) \cdot \sum_{k=0}^{K} \cdot \beta_k \cdot si_k(t) \tag{8.3}$$

Like Table 8.2 shows, the model considers the motivational states of *Rest*, *Play*, and *Socialise*, each tied to a homeostatic process. Table 8.2 also shows the influence of each stimuli on motivational states. The illumination affects the motivation to *Rest*, the user affects the motivation to *Socialise*, and the user and the coloured balls affect the motivation to *Play*. All weights β_k have a value of 0.5 units, excepting the coloured balls stimulus, which had a value β_k of 0.3 units as we considered it as a weaker stimulus for the robot. We chose these weights to balance the effects of the internal deficits and external stimuli on the motivations, bearing in mind our aims at creating a pet-like companion robot. On each time step, the dominant motivation is calculated using a winner-take-all strategy [325]. Motivations can only be a candidate to become dominant if they are above 20 units.

Motivation	**Homeostatic process**	**Related stimuli**	β_k
Rest	Energy	Illumination	0.5
Play	Entertained	User	0.5
		Coloured balls	0.3
Socialise	Social	User	0.5

Table 8.2: Relationships between the homeostatically controlled processes of the robot, its motivational states, and the stimuli involved in the modulation of the motivational intensities.

8.4. Robot behaviours

The deficits in the homeostatic process can only be reduced by executing certain behaviours. Besides, some behaviours can only be executed if the user is in a specific state concerning the robot. In our approach, the user can be in three different states:

- User not detected (User absent). The perception system of the robot does not detect the user.
- User detected but far from the robot (User far). The perception system of the robot detects the robot more than 0.5m far.
- User detected and near the robot (User near). The perception system of the robot detects the user closer than 0.5 meters.

Table 8.3 shows the robot behaviours, their type (consummatory or appetitive), their related dominant motivation, their effects on the homeostatic processes, the state of the user that is necessary for their execution, and a brief description of their purpose in the robot. In our approach, the robot executes behaviours to reduce its physiological deficits and maintain good well-being. In particular situations, the robot has to obtain or perceive specific resources to execute some behaviours, such as perceiving the user near to interact with them. For this reason, MiRo has appetitive and consummatory behaviours, and we consider different user states. On the one hand, appetitive behaviours ease the robot to obtain or perceive a resource in the environment but do not reduce an internal need (e.g. look for the user). On the other hand, consummatory behaviours allow the robot to reduce internal deficits (e.g. talking with the user reduces the social need). If a behaviour is active and influences a homeostatic variable, the corresponding decay rate δ_i of the homeostatic variable defined in Table 8.1 is not applied.

Behaviour	Type	Related Motivations	Effects	User presence prerequisite	Definition
Wander	Appetitive	No motivation	Energy: −0.5	None	The robot moves around in the environment seeking new stimuli
Find user	Appetitive	Play, Socialise	Energy: −0.25	User far or near	MiRo looks for the user inside the enclosure
Approach user	Appetitive	Play, Socialise	Energy: −0.4	User far	Once found, the robot approaches the user to start interacting
Sleep	Consummatory	Rest	Energy: +1.0	None	MiRo falls asleep for a while recovering its energy
Play alone	Consummatory	Play, Socialise	Energy: −0.8 Entertained: +1.5 Social: +0.2	User absent	The robot performs animal-like movements simulating it is playing
Play game	Consummatory	Play	Energy: −0.6 Entertained: +3.0 Social: +0.4	User near	The robot plays a game with the user using coloured balls
Request affect	Consummatory	Socialise	Energy: −0.7 Social: +1.0	User near	Once near, the robot request physical affect from the user by attracting his/her attention

Table 8.3: Definition of the behaviours of the robot during the adaptive social setting. Each behaviour is triggered when its related motivational state is dominant. When active, the behaviour causes different effects on the homeostatically controlled processes. Besides, some behaviour requires the presence of specific external stimuli for being active.

8.5. Emotional model

The affective system designed for MiRo considers four emotional states: *Neutral*, *Happiness*, *Sadness*, and *Surprise*. Using a winner-take-all strategy, the robot's affective state is the emotion with the highest intensity among all active ones. The intensity of each emotion depends on a valence-arousal approach inspired from the Circumplex Model of Affect developed by Russell [190].

8.5.1. Valence dimension

In our model, the valence axis represents the robot's pleasantness. The higher the valence value, the higher the robot's comfort and, consequently, the better its affective state. The valence dimension is a signal ranging from 0 to 100 units, which gathers the influence of internal and external stimuli. Consequently, the valence value in each time step depends on the robot's physiological stability (PS) and the quality of social interaction (IQ), like Equation 8.4 represents. Two weights modulate both the PS and the IQ respectively called $\alpha_{internal}$ and $\alpha_{external}$. On the one hand, $\alpha_{internal}$ represents the internal component of valence, representing the internal state of the robot. On the other hand, $\alpha_{external}$ represents the external component of the robot's affective state, as it is a function of the quality of the interaction. As we will explain below, these weights are essential for the robot to present adaptive behaviour. The default value of both weights is 0.5 units and they act as opposites since $\alpha_{internal} = 1 - \alpha_{extenal}$.

$$\text{valence} = \alpha_{internal} \cdot PS + \alpha_{external} \cdot IQ \tag{8.4}$$

- On the one hand, physiological stability (PS) represents the robot's internal well-being. Higher values of physiological stability mean that the robot does not have any deficits in its homeostatically controlled processes. Low values mean that our robot presents significant physiological deficits. Drawing on similar terminology used in works such as [199, 326, 327], the physiological stability (also comfort or well-being are similar terms) of our robot is defined by a signal ranging from 0 to 100 units calculated using Equation 8.5, where N is the number of homeostatic processes.

$$PS = 100 - \frac{1}{N}\sum_{i=1}^{N} d_i \tag{8.5}$$

- On the other hand, the interaction quality (IQ) represents the external component of the valence value. It depends on the temporal evolution of three signals representing how the user socially behaves with the robot. Each signal is obtained from the robot's perception system and represents the user's presence in the arena, the physiological affect provided by the user, and the kind of answers (correct-incorrect-no answer) responded by the user when playing with the robot. Equation 8.6 represents how the IQ signal is calculated,

$$IQ = \sum_{k=1}^{M} w_k \cdot s_k \tag{8.6}$$

  where $M = 3$ is the number of signals and w_k weights which give more or less importance to each signal on the interaction quality. Note that the adaptive mechanism we present in this chapter regulates the value of these weights to maximise the value of the IQ signal and, therefore, the valence value leading to a happier robot. By default, all weights w_k are set to $1/3$, influencing equally to the IQ signal.

As Figure 8.4 shows, the user's presence is a signal representing the user's time in the enclosure. This signal increases by 1.0 units per second if the user is inside the arena. Otherwise, it is reduced by 0.5 units per second. In the case of the physical affect signal, every time the user strokes the robot, its value increases in a number of units equal to the duration of the stroke in seconds. The value of the physical affect signal is maintained during 60 seconds if any other stroke is perceived. After 60 seconds, the signal value decays by a rate of 0.5 units per second if any other stroke is perceived. If a new stroke is perceived, its effect accumulates to the previous value of the physical affect signal. Finally, the answers signal depends on the answers provided by the user when playing with the robot. If the user provides two consecutive correct answers, the value increases in 40 units. If the user responds correctly, but previous answers were incorrect, the signal increases in 20 units. Two consecutive wrong answers diminish the signal in 20 units, while an incorrect answer being previously correct diminishes the signal in 10 units. The worse case arises when the user does not answer (denies playing). If the robot does not receive an answer, a reduction of 20 units is performed, remarking this reduction to 40 units if more than one positive answer arrives. If the robot and user do not play in 60 seconds, the value of this signal decreases by 0.5 units per second. In this setting, all signals (s_k) range from 0 to 100 units, where a value of 100 is a very intense stimuli p

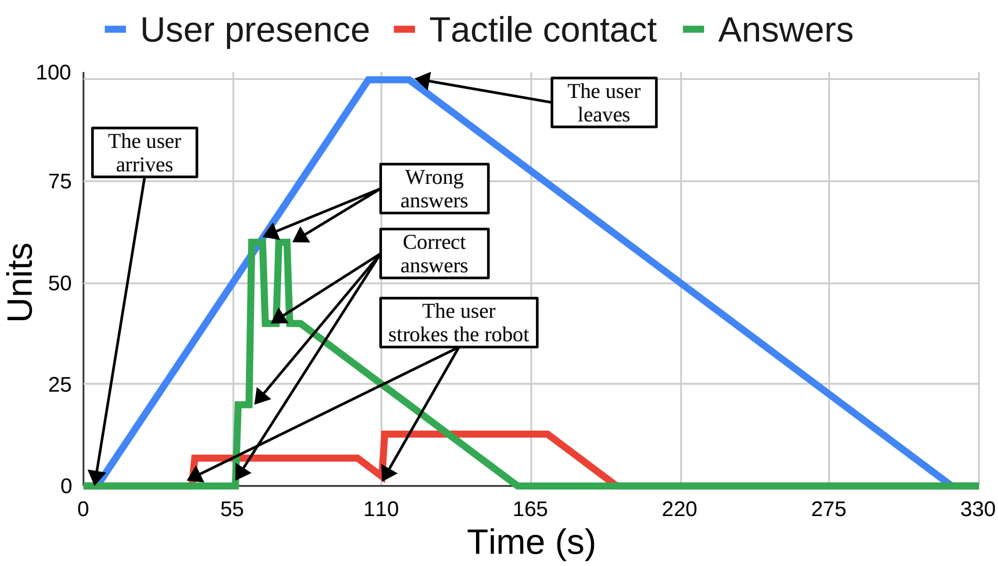


Figure 8.4: Evolution of the external signals representing the evolution of the external stimuli with influence on the IQ.

As defined by Equations 8.4 and 8.6, both the valence value and the IQ signal depend on the values of some weights. The weights $\alpha_{internal}$ and $\alpha_{external}$ respectively represent the importance given to the internal physiological stability (PS) and the quality of interaction (IQ). Both weights are opposites since $\alpha_{internal} = 1 - \alpha_{external}$ and must sum one. Besides, the IQ signal depends on the social interaction between the robot and the user. The weights w_k represent the importance given to each interaction procedure over the external component of the valence axis, having to sum up a value of 1. Note that the values of these weights influence the valence value and, therefore, the robot's affective state. For this reason, the adaptative mechanism we propose in this work aims at regulating the value of these weights ($\alpha_{internal}$ and $\alpha_{external}$) to maximise the robot's internal well-being and minimise its deficits.

8.5.2. Arousal dimension

The arousal dimension represents unexpected sudden changes produced in the environment. The arousal value ranges from 0 to 10 units and is calculated by summing up the variations produced in the signals of the external stimuli (si_k). Thus, the robot will become aroused if the illumination level, the coloured balls, and the user presence signals undergo considerable variations in consecutive time steps.

8.5.3. Robot affect

Our affective model situates four discrete emotions *Neutral*, *Happiness*, *Sadness*, and *Surprise* in the valence-arousal space as shown in Figure 8.5.

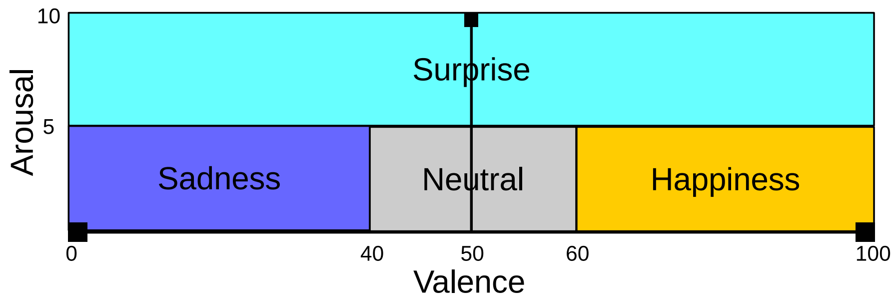


Figure 8.5: MiRo's affective states and their situation in the valence-arousal space.

Emotions have an intensity level (ai_i) from 0 to 100 (excepting Surprise, whose range is from 0 to 120 units), and an activation threshold, which has been set to 20 units for all of them. The output of the affective system is, at each time step, the emotion with the highest level of intensity among all active ones. Each emotion is placed in an interval where they become active in the valence-arousal space. Besides, each emotion has an optimal value inside its interval where its intensity level is maximum. Table 8.4 shows the parameters defining each emotion modelled in MiRo's emotional model.

Affective state	**Intensity ai_i**	**Valence Range**	**Arousal Range**	**Optimal value in valence axis ov_i**
Neutral	[0, 100]	[40, 60]	[0, 5)	50
Happiness	[0, 100]	(60, 100]	[0, 5)	100
Sadness	[0, 100]	[0, 40)	[0, 5)	0
Surprise	[0, 120]	None	[5, 10]	None

Table 8.4: Values of the parameters defining the activation ranges the affective states of the robot.

As Table 8.4 shows, the *Neutral* affective state is situated on the middle of the valence axis, between values $[40, 60]$, *Happiness* on the higher values of valence (interval from $(60, 100]$), and *Sadness* between values from $[0, 40)$. Besides, these affective states require that the arousal value is below 5 units. Inside its activation interval, each affective state has an optimal value (ov_i) used to calculate how fast the intensity ai_i of each affective state increases. In *Neutral*, its optimal value (ov_i) is placed on a valence value of 50, Happiness on 100, and Sadness on 0. Regarding Surprise, it is independent of the valence axis and is elicited at its maximum intensity (120 units) when the arousal value, in a particular time step, is above 5 units. This mechanism allows *Surprise* to become the dominant affective state more easily and independently of the valence value, allowing the robot to react to unexpected stimuli.

The calculation of the intensity level (ai_i) for *Neutral*, *Happiness*, and *Sadness* emotions is carried out following Equation 8.7. If the valence value is inside activation range of the emotion, its intensity (ai_i) increases in ir units. The ir value depends on how close the valence value is to the optimal value of each affective state.

$$ai_i(t) = ai_i(t-1) + ir \tag{8.7}$$

On the other hand, if an affective state is not active because the valence-arousal values are not in the activation range of the affective state, its intensity (ai_i) is reduced by a factor of $dr = 0.2$ units per second, as Equation 8.8 shows.

$$ai_i(t) = ai_i(t-1) - dr \tag{8.8}$$

8.6. MiRo's decision-making

The output provided by the motivational and the modulation exerted by the emotional model respectively influences the robot's decision-making and its expressiveness (Figure 8.6). Thus, the DMS considers the motivational state of the robot to make the appropriate decision about the robot's actions. Then, the expression of the behaviour chosen is modulated depending on its affective state.

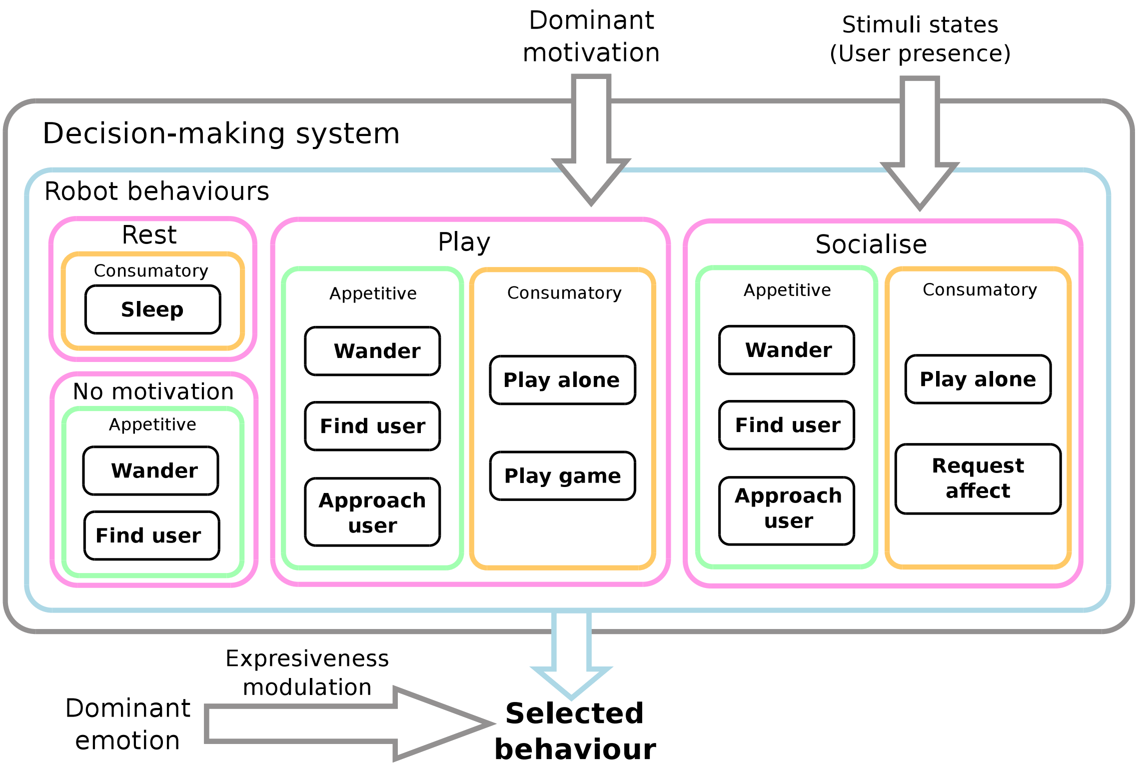


Figure 8.6: The DMS included in the social robot MiRo.

Behaviour selection occurs whenever a previous one finishes or the robot detects a motivational/environmental change. Additionally, changes in the robot's affective state modulate the expressiveness of ongoing behaviours. The decision-making process of the robot works as follows:

1. Behaviours related to the current dominant motivation are selected.
2. The DMS tries to select consummatory behaviours to reduce the deficits of the robot directly. If it is not possible, due to the state of the external stimulus, the system executes appetitive behaviours to obtain new perceptions of stimulus.
3. If more than one candidate behaviour is available, the one with the best benefits for the deficits of the robot is executed. This change is possible because the effects of each behaviour are known by the DMS.
4. The robot adapts its behaviour expression depending on its affective state.

As an example of this process, if the user is absent and the dominant motivation is to *Play*, the robot will execute the *Play alone* behaviour (consummatory). However, if the user is present, the robot will execute a sequence of hierarchical behaviours to *Find the user* (appetitive), *Approach the user* (appetitive) and *Play a game* (consummatory). This order of behaviour selection and execution relies on two aspects.

1. It fosters interaction with users, which is always better than executing solo behaviours.

2. It always finds the behaviours with more considerable reduction rates of the deficits of the robot.

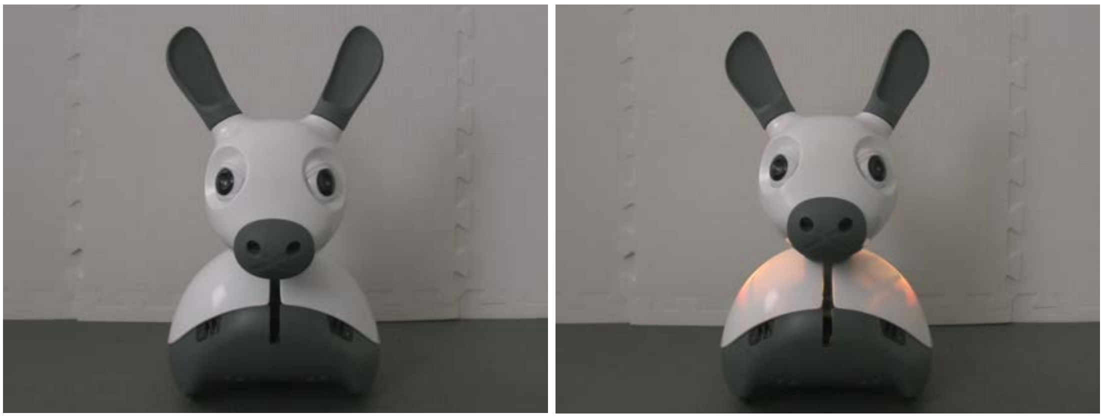
8.7. Expressing affective behaviours

Behaviour modulation throughout affect lets the robot communicate its feelings and mood to the user. The affective state reflects either the robot's internal (physiology) or external (interaction quality) state. Taking inspiration from Arkin's affective robot expression [198] and the work by Castellano et al. [328] about how affective expressions are perceived from human poses and gestures, we opted for varying the amplitude, speed, and frequency of movements of the robot depending on the affective state of the robot. Table 8.5 shows the set of weights used to modulate the amplitude, speed and frequency of MiRo's movements depending on each affective state. These movements comprise the motorised eyelids, ears, head, neck, tail, and wheels of MiRo.

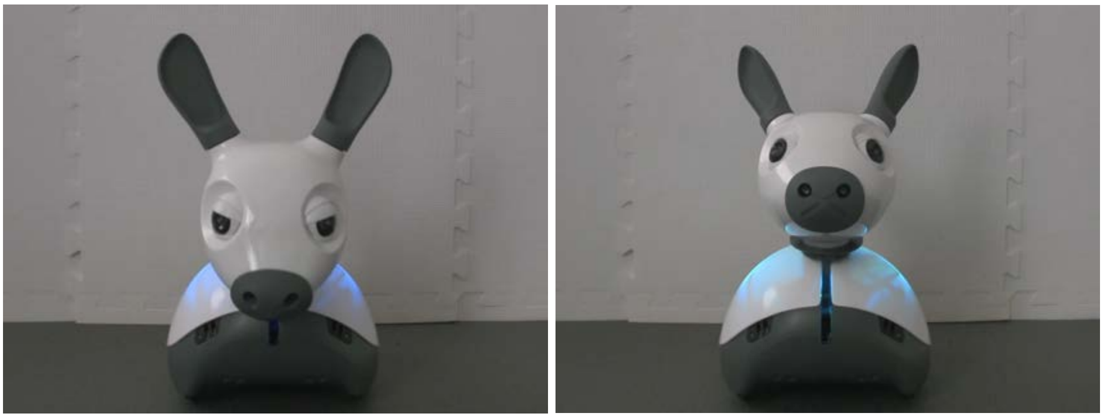
Emotion	Amplitude	Frequency	Speed	Colour
Neutral	0	0	0	Grey
Happiness	0.2	0.2	0.2	Orange
Sadness	−0.2	−0.2	−0.3	Dark blue
Surprise	0.1	0.3	0.3	Light blue

Table 8.5: Values of the weights used for modulating the affective behaviour of the robot during the adaptive scenario.

Affective states have also been related to specific colours as Table 8.5 shows, to aid the user in perceiving the robot's affective state [329, 330]. We have therefore associated each affective state with a colour. Figure 8.7 represents the expression of the affective state on the robot MiRo, considering its body posture and its associated colour.



(a) Neutral (b) Happiness



(c) Sadness (d) Surprise

Figure 8.7: Emotional expressions representing each of the affective states on MiRo.

8.8. Aim of the affective adaptive mechanism

This work presents how regulating the robot's affective state fosters interaction with the user. The adaptive process promotes the interaction by leading the robot to have a happier affective state. As mentioned before, the affective state relies on a valence value dependent on the physiological stability of the robot and the interaction quality with the user, respectively weighted by $\alpha_{internal}$ and $\alpha_{external}$ (Equation 8.4), with the sum of the two weights being 1.

In the adaptive setting, initially, both weights have a value of 0.5. They undergo online variations depending on the value of the physiological signal (PS) and the quality of the social interaction (IQ). If the PS presents a value higher than the IQ, the weight $\alpha_{internal}$ is increased (and therefore $\alpha_{external}$ is decreased) by an adaptive rate equal to 0.01 units. Conversely, if the IQ value is more significant than the PS value, the rate is inversely applied to both weights. This mechanism allows the robot to maximise its valence, driving it to present a happier affective state to engage users in the interaction.

Like the valence value, the IQ is maximised by regulating the weights w_k. Initially, all signals (user presence, answers provided when playing, and physical affect) equally contribute to the IQ (Equation 8.6, all the weights w_k have a value of $1/3$). However, in the adaptive scenario, if the value of a signal is above the others, its weight is increased by a rate of 0.01 units and the others are reduced by a rate of 0.005 units. Thus, the weights of all signals sum up to 1. The adaptive rate defines how fast the weights maximise the valence signal. However, high adaptive rates can produce undesired effects by not perceiving subtle variations in the signals. For this reason, we set the adaptive rate to 0.01 units after assessing which value had the best outcomes in our system.

8.9. Experimental setup and evaluation

The experiment we have designed to assess the performance of our system consists of a simulation phase and real (human)experimenter-robot interactions. In both phases, the robot's behaviour and the experimenter's behaviour were predefined by different profiles. These profiles determine how the experimenter socially interact with the robot and how the robot affectively perceives this interaction. We defined 5 robot profiles and 7 user profiles (described in Sections 8.9.3 and 8.9.2), which led to $7 \times 5 = 35$ different interaction situations for the evaluation. We developed the real experiments and the simulator using ROS [331]. It is worth mentioning that the robot profiles only work in the non-adaptive scenario, as in the adaptive scenario, the robot profile dynamically adapts to the situation that the robot is experiencing.

To assess the robot behaviour, we measured the percentage of time that each motivation and emotion became active when combining a specific user profile with a different robot profile. These metrics aimed to perceive changes in the robot's motivational and emotional states depending on the interaction process. Besides, we used the valence signal to analyse how good the robot's internal state is since the valence value considers both internal and external information.

Simulated trials were performed first to approximate how the system works. It also allowed us to perform many more interactions as simulated experiments are much faster than real ones. Each simulated trial lasted 600 minutes. In this scenario, the architecture was tested during a HRI scenario where both agents behaved following a particular profile. In a second stage, real HRIs were performed, reducing the number of user and robot profile combinations to 3×3, in order to evaluate if results obtained in real domains resemble simulated ones. Each real trial lasted 20 minutes. In the real environment, we carry out two experiments. We combined three user profiles with three robot profiles in the first one. Then, we replaced the robot profiles with the adaptive setting in the second experiment. All simulated and real trials were run once per robot-user profile combination. As mentioned above, including an affective regulatory mechanism in our robot made a more happy robot to engage the users and promote the interaction. This modulation will drive our robot to be emotionally better, fostering positive elicitation, even if the interaction with the experimenter is infrequent or poor (the experiment does not pay attention to the robot or does not provide it with enough positive affect). We hypothesised that by expressing positive feelings to the users that present weak social interaction, the robot could encourage them to interact more often and provide the robot with more affect, engaging the users with the robot.

8.9.1. Scenario

The environment where real interactions took place was a 2m radius area delimited by 0.5m high fences, as depicted in Figure 8.8, where the robot freely moves. When the experimenter is out of the enclosure, the robot's goal is oriented to reduce its deficits, while when the experimenter gets in the arena, becoming perceivable by the robot, the robot's goal will be to reduce its deficits but not to foster social interaction. Recall that social interactions, in this experiment, yield faster reductions of the deficits of the robot. Both agents behave according to a specific profile predefined at the beginning of the trial. In user profiles where the experimenter is engaged with the robot, s(he) will typically provide affective comfort and play with the robot, but for disengaged user profiles, the robot will be the one that takes the initiative.

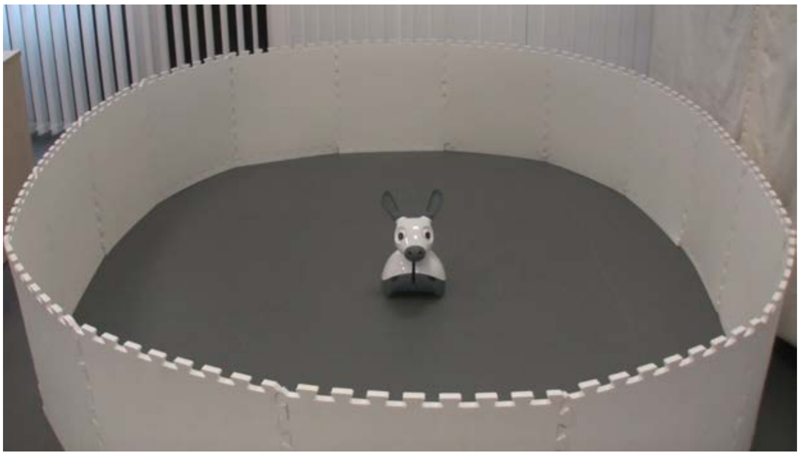


Figure 8.8: The scenario where HRI took place.

8.9.2. User profiles

Table 8.6 shows the features of the user profiles regarding different aspects of the interaction with the robot: the time the user is in and out of the enclosure, the frequency they provide physical affect, and the answers provided when playing together. The following list shows that user profiles can be classified depending on whether the user is engaged or disengaged.

- Fully engaged (FEng): The user is engaged in all interaction modalities.
- Touch engaged (TEng): The user is moderately engaged in all modalities, especially touch contact.

- Game engaged (GEng): The user is moderately engaged in all modalities, especially playing the game.
- Neutral user (NU): A user moderately engaged in all modalities.
- Disengaged with touch (DwT): The user is disengaged excepting in the touch modality.
- Disengaged with game (DwG): The user is disengaged excepting the game modality.
- Fully disengaged (FDis): The user is not engaged in any interaction modality.

User profile	Frequency of physical affect	Time user is in and out the arena	Answers provided when playing
Fully engaged (FEng)	10s	300s inside 60s outside	Correct: 90% Incorrect: 10% No answer: 0%
Touch engaged (TEng)	10s	300s inside 60s outside	Correct: 0.05% Incorrect: 0.95% No answer: 0%
Game engaged (GEng)	60s	300s inside 60s outside	Correct: 90% Incorrect: 10% No answers: 0%
Neutral user (NU)	60s	300s inside 300s outside	Correct: 45% Incorrect: 45% No answers: 10%
Disengaged with touch (DwT)	10s	60s inside 300s outside	Correct: 0.2% Incorrect: 4.8% No answers: 95%
Disengaged with game (DwG)	60s	60s inside 300s outside	Correct: 10% Incorrect: 40% No answers: 50%
Fully disengaged (FDis)	60s	60s inside 300s outside	Correct: 0.2% Incorrect: 4.8% No answers: 95%

Table 8.6: Values of the parameters defining the user profiles used in the evaluation of the adaptive system.

8.9.3. Robot profiles

The robot profiles only work in the non-adaptive scenario by representing how much the valence value in the robot's affective state depends on its physiological stability (PS) and the quality of interaction (IQ). These robot profiles define different values for the weights $\alpha_{external}$ and $\alpha_{external}$, the two parameters that control the importance of the internal state

(PS) and quality of interaction (IQ) on the valence value as defined by Equation 8.4. In the adaptive setting, instead of using fixed $\alpha_{external}$ and $\alpha_{external}$ values, they are dynamically adapted depending on the robot's internal and external situation. Table 8.7 shows the robot profiles we have designed for our non-adaptive experiment. The robot's profiles are:

- Non-dependent on the interaction (nD): The robot's affective state only depends on the robot's internal state (PS) and not on the interaction quality (IQ).
- Mostly non-dependent on the interaction (MnD): The robot's affective state mostly depends on its internal state (PS) and some on the interaction quality (IQ).
- Neutral robot: The robot's affective state equally depends on its internal state (PS) and the interaction quality (IQ).
- Mostly dependent on the interaction (MD): The robot's affective state mostly depends on the interaction quality (IQ).
- Fully dependent on the interaction (FD): The robot's affective state only depends on the interaction quality (IQ).

Robot profile	$\alpha_{\textbf{internal}}$	$\alpha_{\textbf{external}}$
Affective state non-Dependent on the interaction (nD)	1	0
Affective state Mostly non-Dependent on the interaction (MnD)	0.75	0.25
Affective state equally depends on internal and external factors (NR)	0.5	0.5
Affective state Mostly dependent on the interaction (MD)	0.25	0.75
Affective state Fully dependent on the interaction (FD)	0	1

Table 8.7: Values of the parameters defining the robot profiles used during the evaluation of the adaptive system.

8.10. Results

This section presents the simulated and real human(experimenter)-robot interactions results. We divide the results into four sections:

- Simulated experiments without the adaptive mechanism.
- Simulated results with adaptive behaviour.
- Real HRIs without the adaptive system.

- Real HRIs with the adaptive system.

We present the outcomes focusing on the motivational and affective state of the robot and the evolution of the valence component of emotion. Additionally, the evolution of the weights $\alpha_{internal}$ and $\alpha_{external}$ is included in the adaptive settings to represent how the robot adapts its behaviour depending on the situation that it is experiencing.

8.10.1. Non-adaptive robot in simulated domains

Figure 8.9 shows the dominant motivational state of the robot during simulated interactions. The results express the percentage that each dominant motivation was active during the experiment. The adaptive mechanism was not active in this setting, and all weights presented fixed values. The most evident pattern we can observe is how the robot's state is highly influenced by how the experimenter behaves, and the activation percentages of each motivation present a very similar pattern for all robot profiles. This similarity in all robot profiles is a consequence of the independence of the robot profiles (which vary only in how the robot's affective state is calculated) on the motivational state of the robot. The most interesting case is related to the motivation to *Socialise*. In profiles where the user is engaged with the robot, Socialise motivation is not elicited as the user periodically affords its deficit.

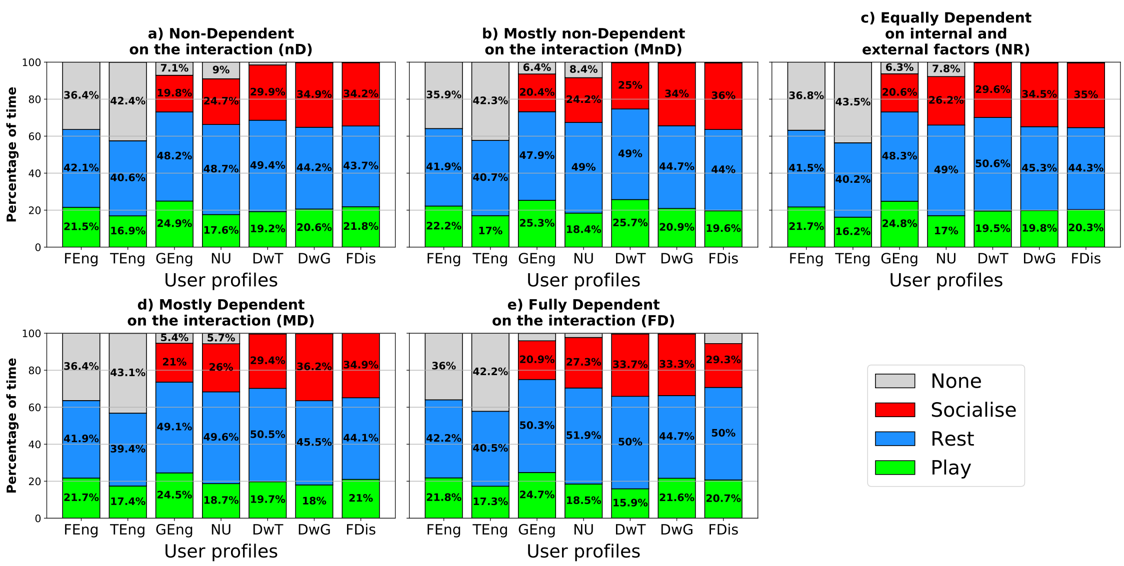


Figure 8.9: Percentage of activation of the robot's motivational states during the non-adaptive trials in simulation. Each graph shows the combination of a specific user profile with all the robot profiles.

Nonetheless, when the user is disengaged (DwT, DwG, and FDis), Socialise becomes active for more time as the robot requests the user to satisfy its social need. It is also noteworthy that the *None* motivational state is not elicited in profiles where the interaction is weak, as the robot expends its time requesting affect instead of being relaxed. The *Play* motivation maintains its percentage of activation in all situations (around 20%), as the reduction of the Entertained deficit can be made autonomously by the robot by playing alone. Nonetheless, in two of the robot profiles, a noteworthy scenario occurs. In Touch engaged (TEng) user profiles, *Play* motivation has valued a little bit lower than in the other profiles (around 17%), as the user satisfies the *Social* need of the robot very often and it can focus on playing with the user as soon as the Play motivation becomes dominant. However, if the user profile is Game engaged(GEng), an opposite effect appears. The activation percentage of the *Play* motivation reaches values around 25% probably because the robot receives much less physical affect than in Fully engaged (FEng) and Touch engaged (TEng) user profiles, so it expends more time requesting affect, driving the *Play* motivation to increase high levels. On the other hand, considering the *Rest* motivation, it is regularly activated to decrease the fatigue of the robot. However, note that the motivation to rest tends to increase its activation percentage for user profiles where the interaction is weak (from Neutral user (NU) to Fully disengaged (FDis)), as it gets more easily tired after interacting with the user.

Figure 8.10 shows the activation percentages of each affective state of the robot on each user and robot profile combination. Substantial differences can be found between user profiles where the experimenter is engaged (profiles FEng, TEng, GEng, and NU) and user profiles where the user is disengaged (profiles DwT, DwG, and FDis) with the robot. It is showed how as we move from robot profiles independent of the interaction (nD, MnD, and NR profiles) to robot profiles that most depends on the interaction quality (MD and FD), the robot's affective state notably worsens. Something worth noting is that in the nD robot profile, Figure 8.10 shows that without the user's assistance, the robot is not able to manage its homeostatically-controlled processes well enough to move its affective state from the Sadness affective state for more than a short time. Conversely, with the user's participation, in all profiles except Disengaged with game (DwG), the robot can manage its homeostatically-controlled processes well enough to keep its affective state at Happiness for most of the time. Examining the graphs in more detail, in engaged user profiles (FEng, TEng, GEng, and NU), when the robot's profile does not depend on the interaction (nD and MnD), Happiness is elicited during most of the time. However, the percentage of activation of Happiness drops when the robot's affective state mostly depends on the interaction quality. This effect is even more notable in disengaged user profiles (DwT, DWG, and FDis). When combined with a robot profile non-dependent on the interaction (nD, MnD), *Happiness* and *Neutral* affective states are elicited during short periods, being Sadness dominant most of the time. However, when the robot affective state entirely depends on the interaction, *Sadness* is triggered during all the trial disappearing residual activation periods of *Happiness* and *Neutral* affective states that were noticeable in robot profiles non-dependent on the interaction quality. The affective state depends on both the deficits of the robot and the interaction quality, being the physiological stability much more stable than the IQ signal. Examples of this effect are when the interaction is impossible (the user is out of the arena). During these periods, the quality of interaction rapidly drops, so robot profiles dependent on the interaction will accuse these effects to a greater extent. Thus, it is clear that as the user becomes more disengaged with the robot, the robot is sad for more time. Additionally, a *Neutral* affective state arises, especially in more neutral user profiles (TEng, GEng, and NU), as the affective state is more unstable and dithers between *Happiness* and *Sadness*. Finally, Surprise presents very low activation percentages in all combinations, as a high level of arousal only happens when the user gets in the arena or strokes the robot after a long time.

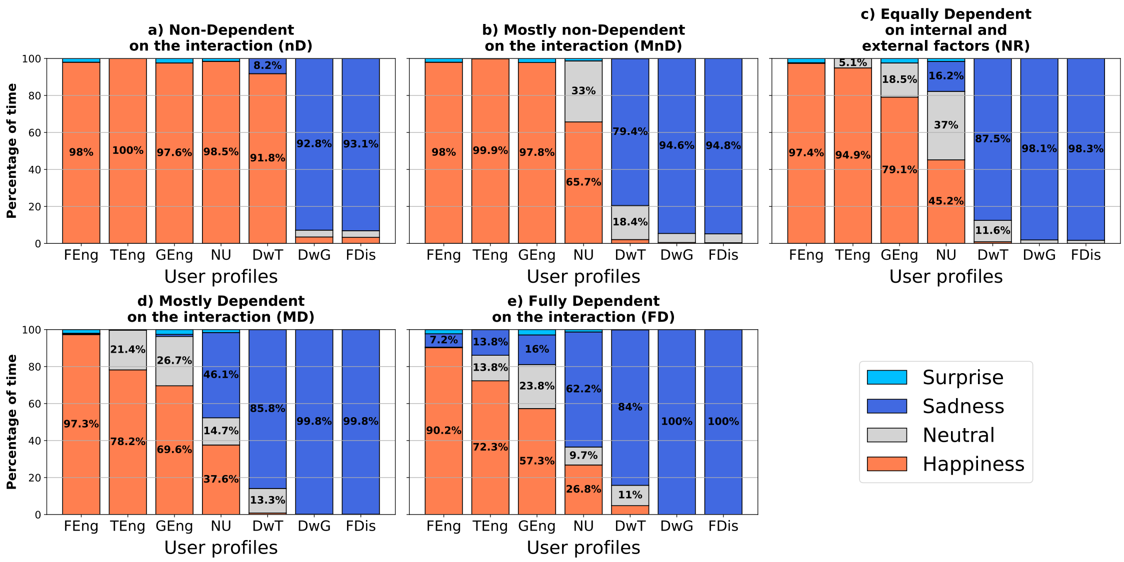


Figure 8.10: Percentage of activation of the robot's emotional states during the non-adaptive trials in simulation. Each graph shows the combination of a specific user profile with all the robot profiles.

Figure 8.11 shows the evolution of the valence signal that strongly influences the development of the affective state of the robot. It reinforces our hypothesis about effect of the affective state on the behaviour of the robot and the experimenter. As the robot profile depends more on the interaction, the valence component variates more. It supports the previous idea that the IQ signal is more unstable than the physiological one. On the other hand, focusing on the user profiles, in disengaged profiles, it is noteworthy how the emotional state of the robot is good at the beginning of the trial but rapidly gets very bad. On the other hand, in engaged user profiles, two different tendencies can be observed: when the robot profiles do not depend on the interaction, the affective state of the robot is excellent, but as the quality of interaction more influences the robot's affective state, the affective state present more considerable variations, especially in user profiles such as Touch engaged (TEng), Game engaged (GEng), and Neutral user *NU*, where the interaction lacks in any of the interaction procedures between both agents. Note that the evolution of the valence on each robot-user interaction represents the percentages of activation previously presented in Figure 8.10, as the affective states of the robot are dependent on its valence and arousal.

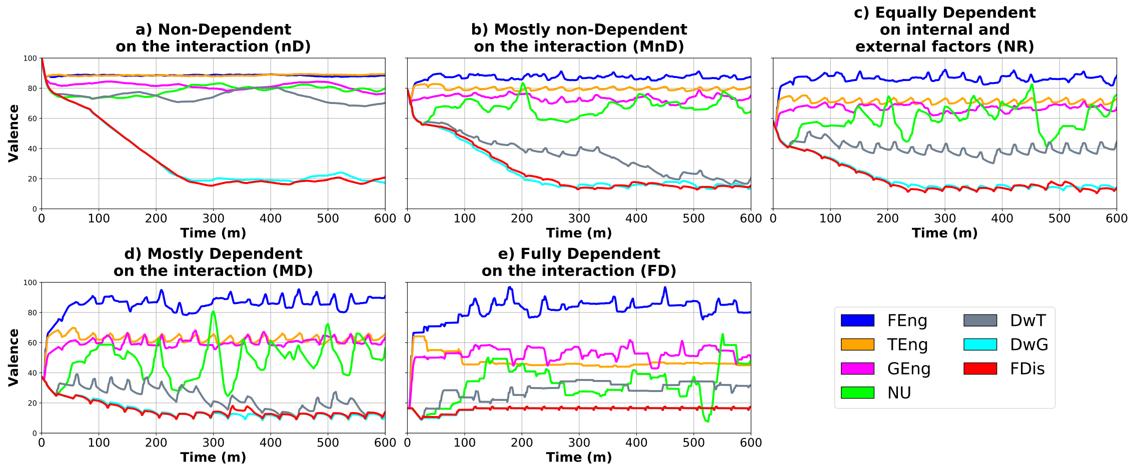
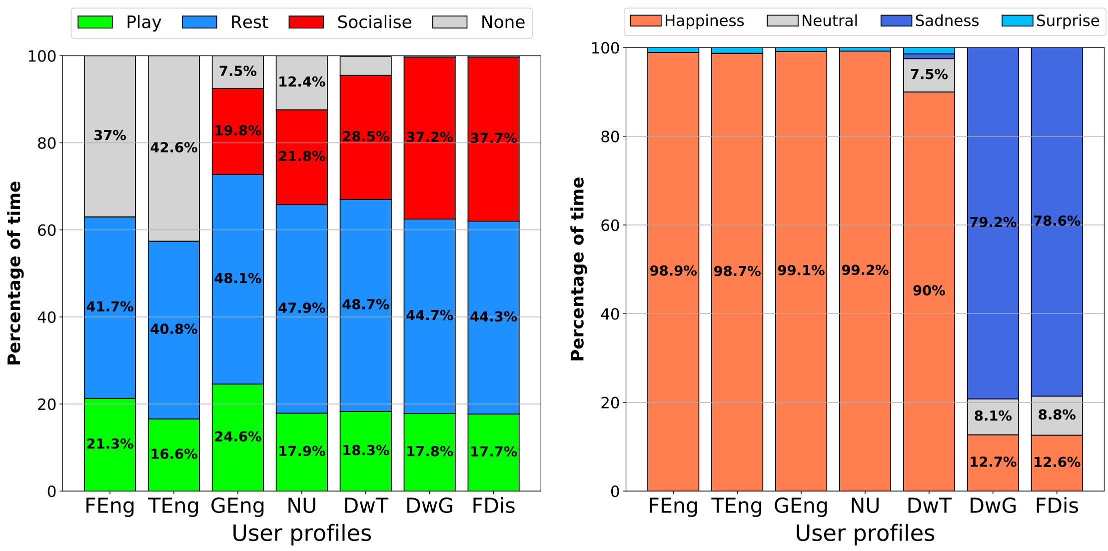


Figure 8.11: Evolution of the valence of the robot during the non-adaptive experiment in simulated trials. Each graph shows the combination of a specific user profile with all the robot profiles.

8.10.2. Adaptive robot in simulated domains

The adaptive mechanism unifies all the robot profiles into a single profile able to regulate the impact of the physiological state (PS) and the quality of interaction (IQ) on the valence component of affect. As can be seen in Figure 8.12a, the activation of each motivational state considering the adaptive model is very similar to the tendency exhibited by the cases represented in the non-adaptive setting (Figure 8.9). The motivational state is not affected by the affective adaptive system and, therefore, no important changes appears for the non-adaptive simulated robot. The impact of the adaptive system is more notorious when the user is moderately engaged with the robot (TEng, GEng, NU, and DwT). The user provides enough interaction in these cases, allowing the robot to maintain a good internal state. However, in the user profiles where the user is strongly disengaged (DwG and FDis), the affective adaptation does not benefit the robot's mood since the PS and the IQ present lower values. The previous ideas are supported by Figure 8.13, showing the valence component of affect is high most of the time for engaged profiles, gets improved in moderately engaged profiles, and is extremely low for disengaged profiles. The evolution of the weight $\alpha_{internal}$ in Figure 8.14 also shows that when the user is engaged with the robot, the valence component of affect is mainly influenced by the IQ signal, but as the IQ worsens, the $\alpha_{internal}$ gets notably increased giving more importance to the PS on the valence component of affect. In disengaged user profiles, the interaction between both agents does not exist or is very weak, leading the adaptive mechanism to maximise the weight of the PS as it presents a better evolution than the IQ.



(a) Motivational state.

(b) Affective state.

Figure 8.12: Percentages of activation of the motivational (left) and emotional (right) states of the robot during the adaptive experiment in simulated domains. Each graph shows the combination of the adaptive robot profile with the different user profiles.

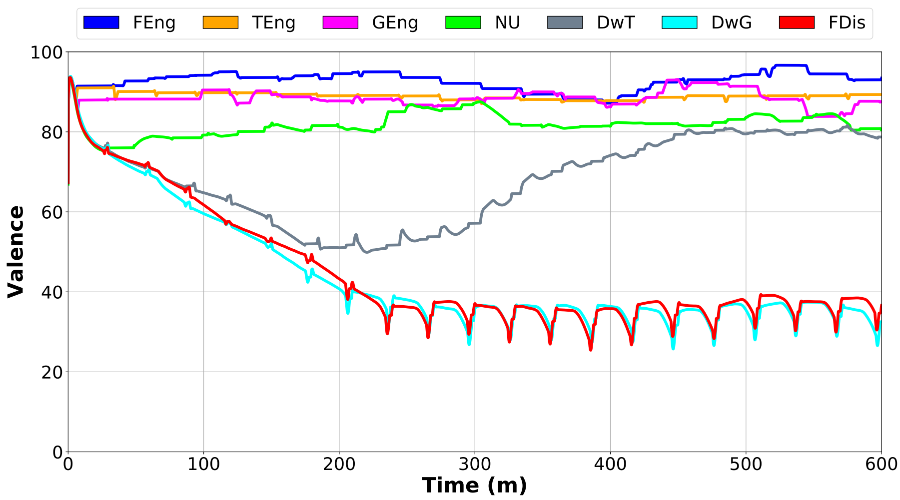


Figure 8.13: Evolution of the valence of the robot during the adaptive scenario in simulation. Each graph shows the combination of the adaptive robot profile with the different user profiles.

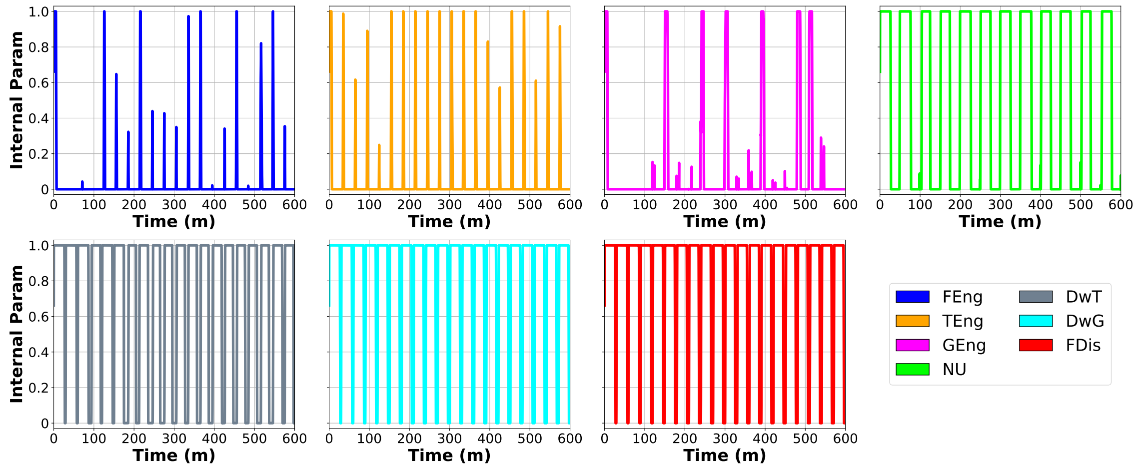


Figure 8.14: Adaptive weights of the internal parameter $\alpha_{internal}$ during the adaptive simulated trials. Note that the value of $\alpha_{external}$ is $1 - \alpha_{internal}$. Also note that each graph shows the evolution of $\alpha_{internal}$ when combining the adaptive robot profile with the different user profiles.

8.10.3. Non-adaptive robot in real HRIs

After simulating the system's operation, the architecture was implemented on the social robot MiRo. These results are presented to support the ones obtained by simulated interaction, as hypothetically, both should produce similar patterns in the robot's behaviour and affective state. Note that in a real experiment, only the user profile Fully engaged (FEng), Neutral user (NU), and Fully disengaged (FDis), and the robot profiles Mostly non-dependent on the interaction (MnD), Neutral robot (NR), and Mostly dependent on the interaction (MD) to reduce the combination space. The duration of each trial was also reduced to 20 minutes. In these settings, the adaptive mechanism was not included, and the weights were set to fixed values as defined in Tables 8.6 and 8.7.

Figure 8.15 shows the motivational state of MiRo in real interactions. In this case, the hypothesis about the similarity with simulated interactions is noticeable at first glance. Play and Rest motivations are active most of the time on each trial, presenting similar activation percentages to simulated runs. However, the robot normally plays more with engaged users, probably because the robot's Social deficit is satisfied more often than in disengaged user profiles. Considering the motivation to Rest, its activation percentages vary between 30% and 50%, being maximum in Neutral user (NU) and Fully disengaged (FDis) user profiles as the robot gets tired more easily. Considering the motivation to Socialise, its activation increases for disengaged user profiles (DwT, DwG, and FDis). For simulated results, None motivation gains importance as it is active during more time, maybe due to its activation by default at the beginning of the trial.

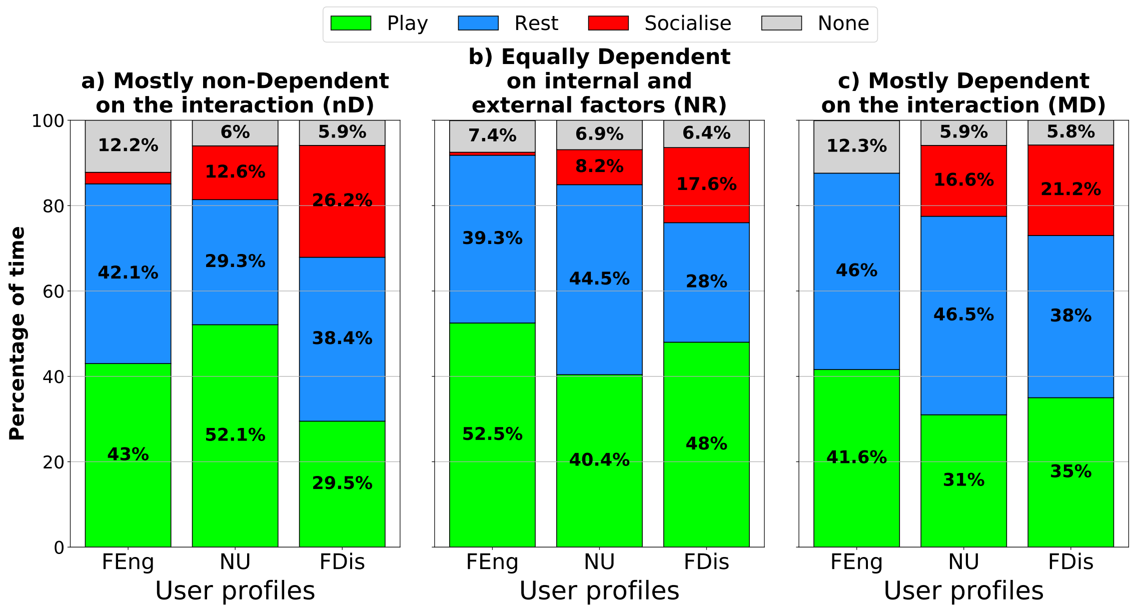


Figure 8.15: Percentages of activation of the motivational states of the robot during the non-adaptive real experiment. Each graph represents the results of combining a robot profile with all the user profiles.

Affective states in MiRo are also activated with similar percentages in real and simulated trials (Figure 8.16), supporting the previous hypothesis. Happiness is dominant when the user is Fully engaged (FEng), independently of the profile set in the robot. Most of the time, a Neutral user (NU) led to a sad and affectively *Neutral* robot, being this effect even more predominant when the user is Fully disengaged (FDis) in all modalities. *Surprise* is elicited with percentages around 5% for engaged users. However, its activation decays when the user is fully disengaged due to the lack of interaction.

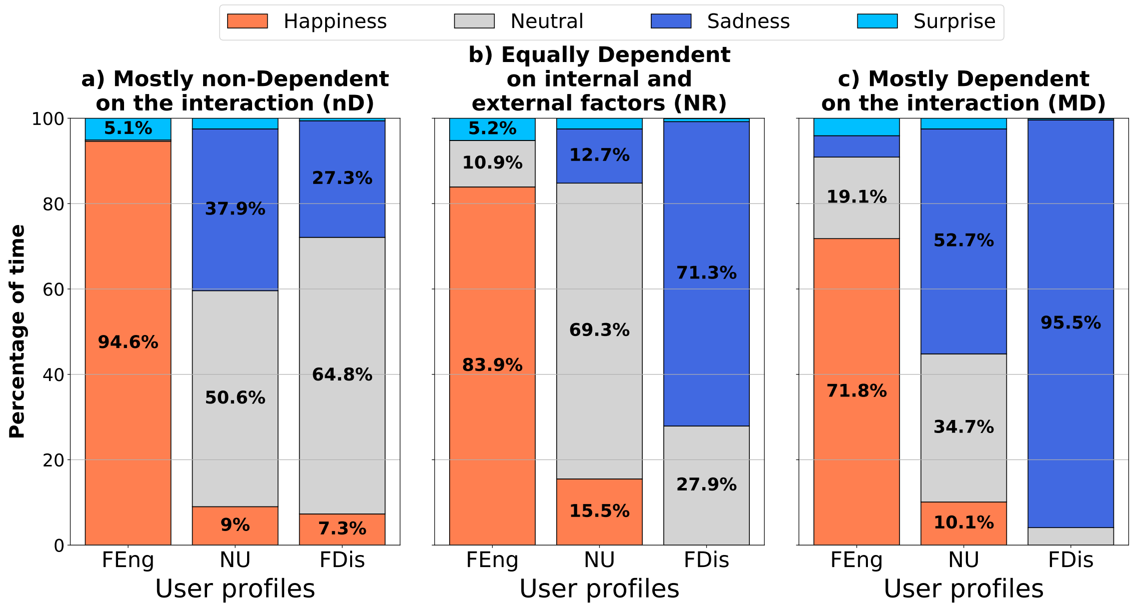


Figure 8.16: Percentages of activation of the emotional states of the robot during the non-adaptive real experiment. Each graph represents the results of combining a robot profile with all the user profiles.

Looking at Figure 8.17, it is possible to perceive how, independently of the robot profile, an engaged user maintains the robot in a stable and very positive affective state. However, the robot's affective state rapidly drops if the user is moderately or fully disengaged. In the case of a Neutral user (NU), it is represented how the robot's affective state tends to be in a neutral state. It is important to be aware that while in the motivational state, the robot is comfortable when its dominant motivation is None. In the case of the affective state, Happiness is the optimal one. Having this in mind, the results demonstrate how engaged user profiles make the robot physiologically comfortable and emotionally happy, but when the user is disengaged, the robot's state is deficient in both ways. For this reason, the following Section 8.10.4 provides the results obtained by replacing the robot profiles with an emotionally adaptive robot that dynamically regulates its affective state to become more positive when the interaction quality is not good, conveying a positive feeling to the user.

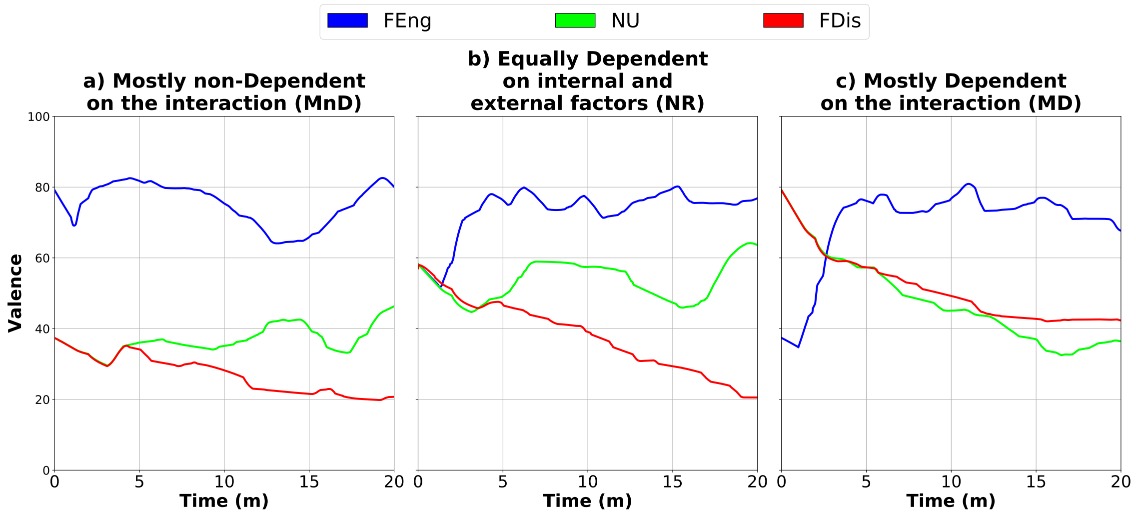
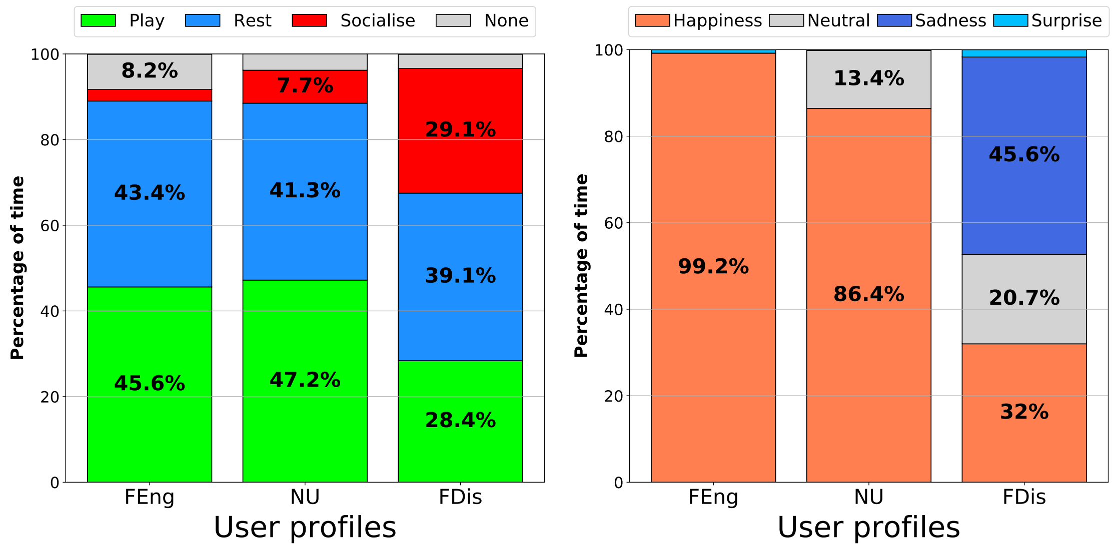


Figure 8.17: Evolution of the valence value during the non-adaptive real experiment. Each graph represents the results of combining a robot profile with all the user profiles.

8.10.4. Adaptive robot in real human-robot interaction

Results in Figure 8.18 (left) demonstrate how the motivational state of our robot maintains similar activation percentages that results presented without the affective regulation. This problem is because the physiological state of the robot does not depend on the affective state regulation, as their internal homeostatic processes evolve equally in all situations. Although the emotional regulation of the robot does not affect its motivational state, Figure 8.18 (right) shows how the affective state substantially improves in terms of the elicitation of Happiness, especially in the Neutral user (NU) and Fully disengaged (FDis) robot profiles (whereas in non-adaptive trials was not elicited). Happiness is dominant in practically all the trials when the user is Fully engaged (FEng), and *Sadness* does not appear in the Neutral user *NU* profile. Nonetheless, Happiness has notably increased its activation percentages in the Fully disengaged (FDis) user profile. Sadness still appears during long periods. This issue can be explained by thinking about how adaptation occurs. If interaction with the user is lacking, the robot's affective state will be mostly influenced by its physiological state. Nevertheless, the physiological state is affected by social interaction in terms of physical affection and playing together, so if these deficits do not decrease, the affective state cannot be optimal during some periods. Figure 8.19 supports this idea in the sense that it describes how the valence signal, which is the basis of the affective state of the robot, has been significantly improved in all profiles. However, the figure shows how, in the case of Fully disengaged (FDis) user profile, the tendency is to deteriorate as the trial goes forward. Additionally, the evolution of the weight $\alpha_{internal}$ (Figure 8.20) supports the previous explanation. The $\alpha_{internal}$ weight presents low values when the user is engaged with the robot (FEng) because the IQ is excellent and the physiology of the robot always present some deficit. This weight gains importance when the interaction is lacking, as if the IQ presents shallow values. The adaptive mechanism regulates the weights increasing $\alpha_{internal}$ to strengthen the importance of the PS on the valence component of affect, seeking to maintain a good affective state to maintain user engagement.

To conclude the section, we have described how the robot's affective state improves in most cases, despite when the user is very disengaged and the interaction lacks, the adaptive system cannot attain its goal. Nonetheless, the notable improvement of the robot's affective state drives it to express positive expressions when interaction exists but is weak, encouraging the users to be more engaged with the robot.



(a) Motivational state.

(b) Affective state.

Figure 8.18: Percentages of activation of the motivational (left) and emotional (right) states during the adaptive real experiment. Each graph shows the results of the combination of the adaptive robot profile with the three user profiles considered in the real experiment.

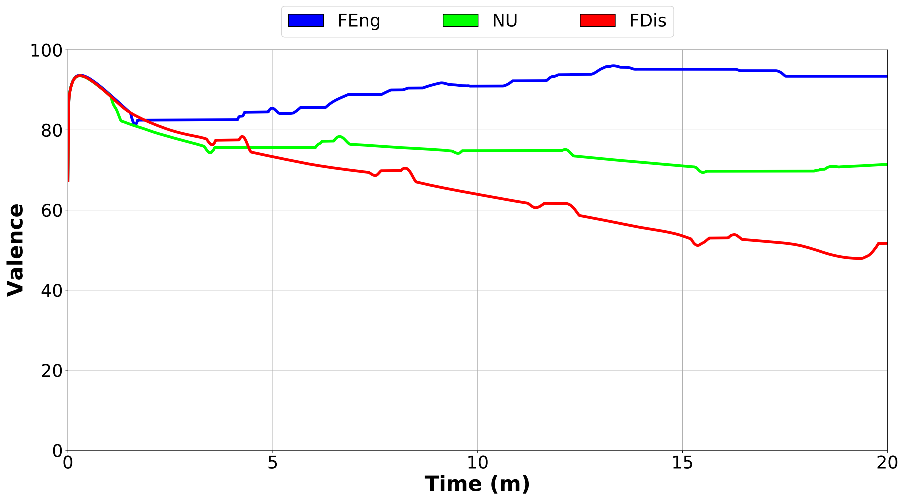


Figure 8.19: Evolution of the valence signal during the real adaptive scenario. The graphs shows the results of combining the adaptive robot profile with the three user profiles considered in the real experiment

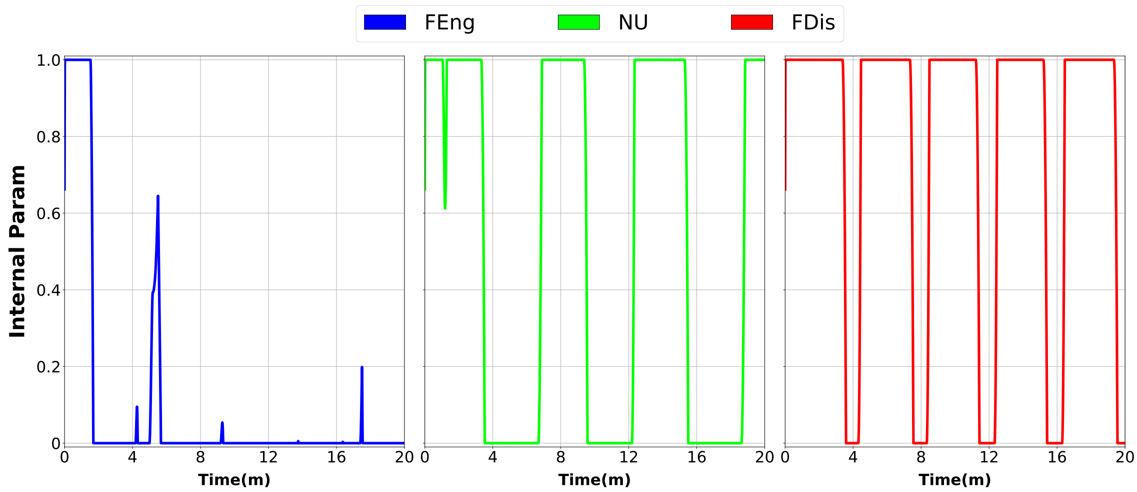


Figure 8.20: Evolution of the adaptive $\alpha_{internal}$ weight during the real adaptive experiment. Note that $\alpha_{external}$ is $1 - \alpha_{internal}$, as both weights are contraries.

8.11. Discussion

In this work, we have presented a robot architecture supported on biologically-inspired motivational and affective concepts used to engage users that may show interaction limitations. The robot's affective state regulation can transmit positive affect to engage the user. The experiments we carried out sought to demonstrate how the behaviour executed by the user highly influences the affective state of our robot. In such a way, we predefined a set of user and robot behavioural profiles, which allowed us to assess the architecture in two stages. First, we evaluated how the combination of user and robot profiles affects the robot's motivation and emotions in the simulation. After, we carried out the same experiments in real HRIs to prove if the robot's state undergoes the same evolution depending on the behaviour of the experimenter. Finally, the adaptive mechanism replaced the robot profiles to improve the robot's affective state, especially when the interaction with the user is weak. By including an adaptive affective state that fosters positive affective states, we want to engage the user in the interaction.

Simulated runs showed how different user profiles affected the robot state and its behaviour. As a clear example, the Socialise motivation triggers more often when the user is disengaged (DwT, DwG, and FDis), as the social need presents a high deficit. Besides, while in engaged user profiles (FEng, TEng, and GEng), the robot feels comfortable with all its deficits satisfied, disengaged user profiles (DwT, DwG, and FDis) leads the robot to be more aroused as it requests affect from the user and therefore gets more tired. Considering the results obtained in real interactions, their similarity to simulated runs is noteworthy.

Focusing on the robot's affective state, it is possible to perceive that its variability depends on its behaviour. Engaged user profiles (FEng, TEng, and GEng) elicit Happiness most of the time, obtaining a very happy and expressive robot. Nevertheless, as we move towards disengaged users (from NU to FDis), the affective state worsens, and Sadness becomes the dominant affective state more often. Regarding Surprise, it is easily elicited when a good interaction is produced. As results show, *Neutral* appears during short periods in user profiles like Touch engaged (TEng), Game engaged *GEng*, Neutral user (NU), and Disengaged with touch (DwT), where the interaction quality is moderate. The valence evolution aids us to understand why different motivations and affective states are elicited, as are the bases of affective elicitation. In disengaged profiles, we can see how the valence component of emotion varies more due to the instability of the interaction. Real interactions without the adaptive mechanism provide very similar results, as activation percentages of each affective state and the evolution of the valence presents the same tendency. The user profiles aim at improving the robot's affective state when including the affective adaptive mechanism. This improvement is especially remarkable in moderately engaged user profiles (GEng, NU, and DwT). However, the notable lack of interaction in some of the disengaged user profiles (DwG and FDis) makes the adaptive system not attain its goal of improving the robot's affective state. Thus, improving its affective state drives our robot to be more cheerful and expressive, encouraging the user to interact with it.

As mentioned before, this work was carried out during a research stay at the University of Hertfordshire. This period supposed a strong source of inspiration and learning for researching how to endow our social robot Mini with adaptation and personalisation to each user. This influence can be perceived along the manuscript but specially in the neuroendocrine *Motivational model* we presented in Chapter 6 and their associated results present in Chapters 10, 11, and 12.

Chapter 9

Evaluating the robot's Preference Learning System

This chapter focuses on describing our Preference Learning (PL) System used for personalising HRIs. This system, introduced in Section 4.2, allows the decision-maker to select those entertainment activities that the user likes the most. The PL System works in three steps.

- **User profiling:** For the PL System to work, the robot has to obtain, using HRI, the user's features for producing an initial estimation of their preferences. Additionally, the profiling includes taking photos to train a facial recognition system so the robot can identify the user in future interactions.
- **Estimating preferences:** Using the user's features retrieved in the previous step, the robot uses a Label Ranking (LR) algorithm to predict their initial preferences.
- **Adjusting the initial estimation:** After the prediction, the robot uses Reinforcement Learning (RL) to adjust, using the information gathered from the interaction, the initial estimation to the real user preferences.

Figure 9.1: Our proposed Preference Learning (PL) System. It consists of three stages: user profiling, estimating their preferences, and adjusting the initial estimation using the interaction.

9.1. User profiling

The user profiling method consists of applying *Active Learning*, a machine learning technique that allows our robot to dynamically recognise and obtain information from the users that interact with the robot [332]. The goal of the profiling stage is to obtain as much information from the user to personalise HRI to them. Our methodology proposes to apply *Active Learning* in two stages of the profiling process:

1. Learning the robot users' faces to dynamically update a facial recognition system.
2. Filling the user profile attributes using HRI.

9.1.1. Profiling method

As we mentioned in Section 7.4.3.3 and Figure 9.2 represents, the dynamics of the method vary depending on whether the robot knows or not the user. Consequently, the *Active Learning* process starts when the visual perception system of the robot detects a user ready to interact seated in front of it. When the user is in the correct position, the perception system informs the DMS whether the robot already knows the user or not, using a facial recognition algorithm developed by Google based on neural networks [333].

Consequently, depending on if the robot knows the user or not, the operation of the *Active Learning* for user profiling and identification works as follows.

- If the robot does not recognise the user, the DMS starts the *User profiling skill* as described in Section 5.2.2.1. Once active, the *User profiling skill* conducts a user profiling method that allows the robot to create from scratch a unique user profile that will be completed using HRI. All user profiles are created and managed by the *Context manager* (see Section 5.2.2.2).

  The profiling method consists of the robot asking different questions to the users. First, the robot asks for their name and surname. Once the robot gathers this information, a unique id is assigned to the user. The profile is stored in the *Context manager*, so the *User profiling skill* communicates with this module to check whether the user profile has been successfully created. Once the essential information has been retrieved, and the profile has been created, the robot takes 5 pictures of the user's face and stores them. The user must be sitting in front of the robot, looking at the front for taking the images. To avoid taking bad-quality images, the robot can guide the positioning of the user giving voice commands and using a head pose estimator [334]. In this scenario, the *Active Learning* takes place on two occasions. First, when the robot obtains the user's name and surname, it actively updates its profile. Second, when the robot obtains pictures of the users' faces for identifying them in the future, the facial recognition algorithm is trained online to include the new user's information.

- If the robot detects a known user, it requests the *Context manager* to load the user's profile. Then, the robot checks if some user profile attributes are unknown or can be updated. If any attribute can be retrieved, the robot will use HRI to continue

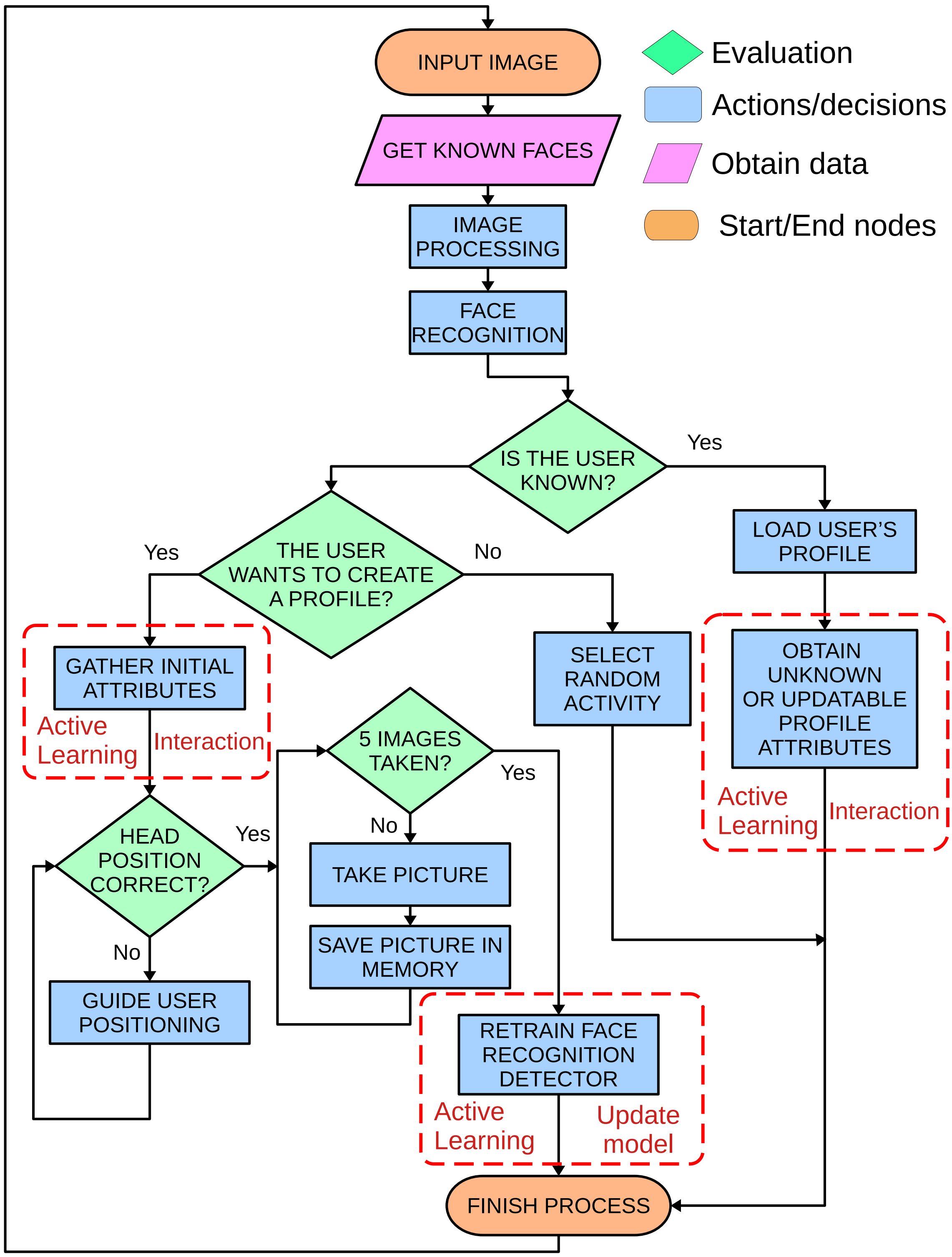


Figure 9.2: Diagram showing the active user identification and profiling method.

gathering more information from the user. The filling of the user profile actively works using priorities and randomness to avoid repetition. As mentioned in Section 5.2.2.2, the robot stores personal, habits, and preferences information about each user. Inside the personal attributes, it is possible to find from general information like the sex, age, or date of birth to more meaningful information such as physical or psychological impairments. The habits section contains information like the frequency of performing physical activity or their interests in music, literature, and other related trends. Finally, the robot stores the user's preferences to the robot's entertainment activities to select their favourite activities and adapt to the interaction. In this scenario, the *Active Learning* takes place when the robot uses HRI to update the user profile with new information, improving the personalisation of its behaviour towards the users' features.

The *Active Learning* is controlled by the DMS, which activates this functionality every time the robot is interacting with people. Thus, once in a while, the DMS activates the *User profiling skill* to continue gathering information from the user. In the first place, the robot focuses on collecting the user's unknown attributes in the profile for completing it as soon as possible. If more than one attribute is unknown, the robot will ask those with a higher priority level. For example, personal attributes are more important than attributes about preferences. If more than one attribute with the same priority is unknown, the robot will select which one it fill randomly. If the profile has all the attributes completed, the robot will hardly ever repeat questions to update some profile attributes with new information. It is important to note that not all the attributes can be updated, as some are fixed (e.g. sex or name).

Every time the robot obtains or updates an attribute of the user profile, the *User profiling skill* sends it to the *Context manager* for its storage in the robot's memory. Then, the *Context manager* updates this parameter in the robot's *Memory*, making it accessible to all the modules and applications in the software architecture. During the interaction, the robot can adapt its behaviour depending on the values of these parameters. Thus, the robot can call users by their name, suggest their favourite applications, or control their proactivity level.

The recognition and profiling process was validated in an experiment where the participants completed the process interacting with different robot profiles and using an online survey. The following sections detail the experiment, its evaluation, and the results.

9.1.2. Experimental setup

The *Active Learning* for user profiling was tested in an experiment where we considered three conditions. We use the social robot Mini [293] for performing such experiments.

- The first condition (Condition 1) consisted of creating the profile using an online survey instead of interacting with the robot. 19 participants (11 women and 8 men) participated in this condition.

- The second condition (Condition 2) aimed at completing the profile interacting with a *dull robot*. A *dull robot* can be defined as an agent who lacks the expressiveness and focuses on its task without pretending to engage the user. 15 participants (6 women and 9 men) participated in this condition.

- Finally, the third condition (Condition 3) consisted in the users creating their user profile interacting with a *cheerful robot*. We define a *cheerful robot* as an agent that attempts to engage the user by being kind and expressive while executing its task. 15 participants worked with this condition (7 women, 8 men).

The primary goal of conducting such an experiment was to test the system's operation and, at the same time, investigate if users consider it more engaging to create their profile interacting with a *cheerful robot* than interacting with a *dull robot* or completing an online survey. It is worthy to stand out the following aspects of the experiment.

1. All conditions had 49 questions about personal information, habits/interests, and users' preferences towards the robot's activities.

2. In the condition that employed an online survey (Condition 1), it was impossible to capture the users' faces. For this reason, this process was moved to the end of the two conditions using the robot (Conditions 2 and 3) to avoid biasing the evaluation.

3. All participants interacted with Mini in one-to-one interactions. These interactions occurred with the user interacting alone with the robot without supervision.

After completing the profile, the participants of all conditions had to fill out an online survey to evaluate the profiling process. It is worth noting that filling the profile using an online survey incorporated the evaluation survey at the end. The following section describes how we compared and statistically evaluated the three conditions of the experiment.

9.1.3. Evaluation

The evaluation of the experiment was performed using the well-known System Usability Scale (SUS) questionnaire [335]. This questionnaire, included in Appendix D.4, measures the system usability from the user's point of view using ten questions and a 5-point Likert scale. Figure 9.3 shows the format of the questions included in the questionnaire. As the target population of the study were Spanish, we used the Spanish version of the questionnaire proposed in [336]. Each question rated 1 to 5, where 1 means strongly disagree, and 5 means strongly agree. The marks of each questionnaire map to a $0 - 100$ scale that represents the usability value of the system. Grades below 50 units mean weak usability. From 50 to 70 the system is correctly usable, from 70 to 80 the usability is good, and above 80 the usability of the system is excellent [335].

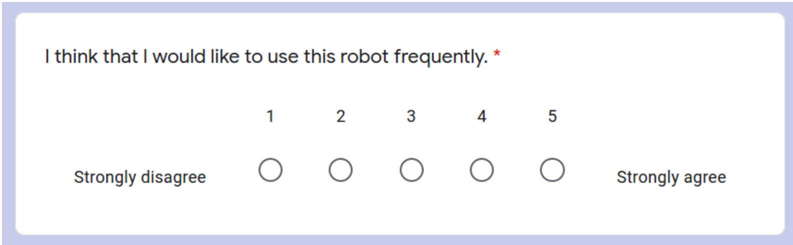


Figure 9.3: Example of the SUS questions asked to the participants for evaluating the system's usability.

Additionally to the SUS questionnaire, we included four ad-hoc questions (included in Appendix D.2) asking the user about how engaging the user profiling process was at different stages of the process. The answers were bound to a 5-point Likert scale where participants had to indicate how engaging the process was from 1 (not engaging at all) to 5 (very engaging). These questions appeared when the 20%, 40%, 80%, and 100% of the profiling progress was completed. Next, we provide the results obtained from the system's evaluation.

9.1.4. Results

Figures 9.4 and 9.5 show the results obtained from the experiment in terms of usability and user engagement considering the three conditions previously stated: filling the user profile using an online survey (Condition 1 in blue), filling the user profile interacting with a *dull robot* (Condition 2 in green), and interacting with a *cheerful robot* (Condition 3 in red).

In terms of usability, as Figure 9.4 shows, the condition where users interacted with a *dull robot* obtained the best average score ($\mu = 83.21, \sigma = 10.45$), followed by the condition where users interacted with a *cheerful robot* ($\mu = 82.14, \sigma = 9.49$), and ending with the condition where the users filled their profile using an online survey ($\mu = 71.05, \sigma = 13.31$). Besides, we explored whether the usability scores depend on the evaluation condition. To carry out such a test, and after proving that data is generally distributed in all conditions, we used a one-way ANOVA study [337].

The ANOVA results showed that the usability scores were affected by the condition used to carry out the profiling. Considering these results, we analysed which conditions produced significant differences between them, carrying out a posthoc Tukey test [338]. The study showed significant differences between Condition 1 (Online survey) and Condition 2 (*dull robot*) ($p = 0.014, \eta^2 = 0.202$) and between Condition 1 (Online survey) and Condition 3 (*cheerful robot*) ($p = 0.027, \eta^2 = 0.202$). However, none statistical difference was obtained between both conditions using the robot ($p = 0.968, \eta^2 = 0.202$).

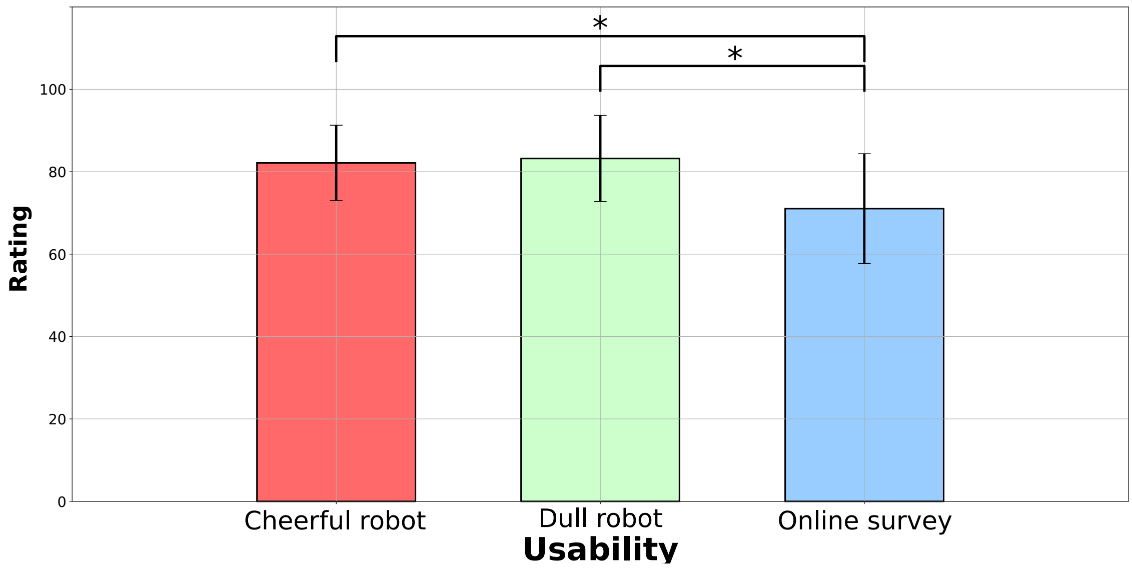


Figure 9.4: Usability ratings for each condition assessed in the user profiling experiment. Significant statistical differences were found between those conditions using the robot for completing the profile and the online survey process. However, no difference was obtained when comparing both conditions using the robot. Significant statistical differences (p-value < 0.05) are denoted with $*$.

Considering the user engagement during the profiling process (Figure 9.5), results show that users found more engaging using the robot than completing the process using an online survey. The participants that completed their profiles using an online survey reported the lowest values in all checkpoints. The condition where the robot behaves cheerfully obtains the best ratings excepting the last checkpoint, where the condition including the *dull robot* gets a score slightly above the condition including a *cheerful robot*. This fact is probably due to the longer lasting interaction required when completing the profile with a *cheerful robot* (Condition 3). Besides, it is worth noting that the engagement decays with time in Condition 3, where the *cheerful robot* is used, contrary to Condition 1, where participants filled their profile using an online survey. Condition 2, where the user interacted with a *dull robot* also shows a decay in the user engagement, despite its scores being more balanced than in the other cases. As in the usability study, we performed a statistical analysis to find significant differences in the user engagement for the three conditions. However, no significant differences were obtained in this analysis.

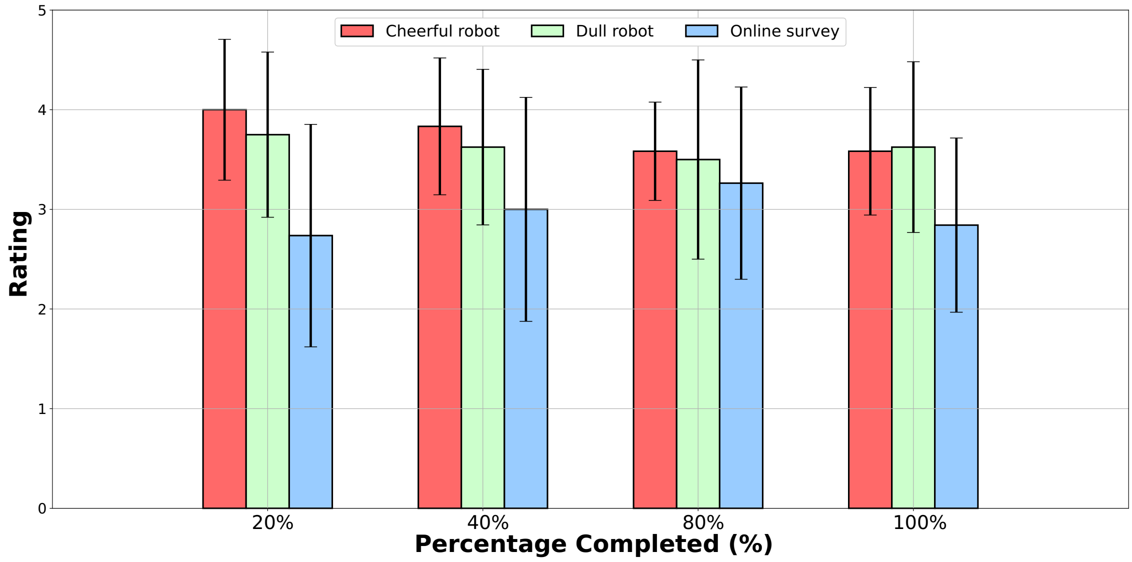


Figure 9.5: User engagement at three different moments of the profiling process.

9.1.5. Discussion

As expected, during the profiling experiment, the usability of Conditions 2 (dull robot) and 3 (cheerful robot) obtained better scores than Condition 1 using an online survey. However, users perceived a *dull robot* subtly more usable than the cheerful one. Considering the scores obtained in the evaluation, we can conclude that the usability of the robot is excellent, while the usability of the online survey is just good. Additionally, the significant differences found in the statistical analysis between both conditions, including the robot and the condition using the online survey, demonstrate that participants believe that completing the process with the robot instead of using an online survey is more usable.

On the other hand, the user's engagement during the experiment showed very different results depending on the evaluation condition. As expected, the engagement is higher in those conditions where the robot participates. In the beginning, users find a *cheerful robot* more engaging than a *dull robot*. Nonetheless, looking at the last bar group (100% completed) in Figure 9.5, users consider the *dull robot* slightly more engaging than the cheerful one. A possible explanation for this issue is the length of the profiling process. Since a *cheerful robot* articulates more phrases to entertain the user, the duration of the process increases. This issue can work against us, causing decay of the engagement in the last step of the process. Finally, considering the condition where the users completed their profile using an online survey, results show that participants were much less engaged than those using the robot. At the beginning of the profiling, the engagement is weak (below 3 from a maximum of 5). This value rises in the middle of the process to 3.2 units. However, they drop again at the end of the experiment. Despite not having a clear explanation for this, the comments provided by the participants who filled the online survey suggest that they found the survey a bit longer and repetitive at the end.

9.2. Estimating the user's preferences

The second step in our PL System comprises the prediction of the preferences using the user features retrieved in the previous step. Following, we describe the methodology, evaluation, experiment, and results obtained at this second step included in our PL System.

9.2.1. Designing the learning environment

The PL scenario aims at providing the social robot Mini with the capacity of selecting those entertaining activities that the user likes the most. Different strategies and experiments were tackled to endow robots with such functionality. Taking inspiration from how recommender systems and websites learn the preferences of customers and estimate the star products, we studied PL algorithms [339] for Label Ranking (LR) classification. As detailed in Section 5.2.3, the robot has many entertainment activities that are classified into three hierarchical levels. These levels organises the activities (level 3) in categories (level 1) and subcategories (level 2), as Figure 9.6 shows. Considering this classification, the LR algorithm should produce, for each particular user, a ranking per category, subcategory, and groups of activities. The rankings indicate the preferences towards the entertaining activities of the robot organised in such three levels based on the user features obtained during the profiling. As will be described in the following section, the PL System predicts the rankings for each new user considering their features and the preferences of similar users whose features and preferences are stored in a database used for machine learning classification. It is worth mentioning that the prediction stage uses LR based on Random Forest, so the estimated rankings grow from associations made from the dataset containing user features and their labelled rankings.

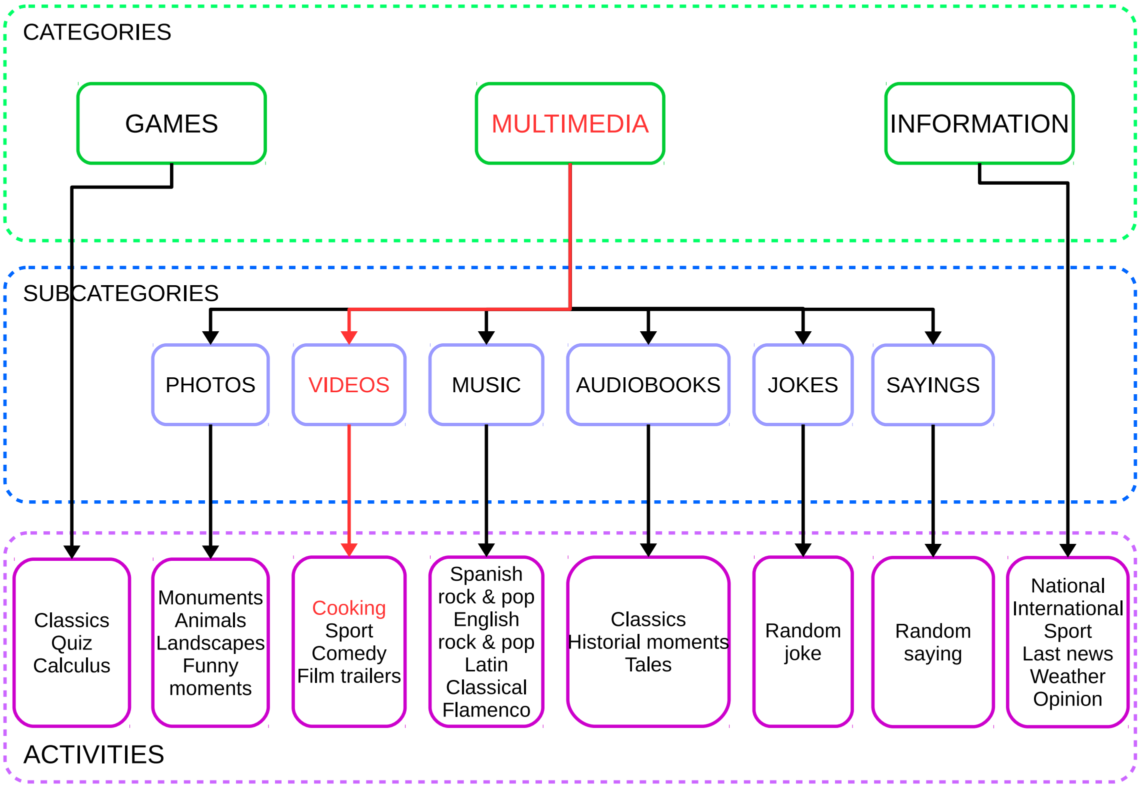


Figure 9.6: Robot entertainment activities organised in hierarchical levels and an example of how the DMS selects an activity following three hierarchical steps. In this case, the selection is the Multimedia category, the Videos subcategory, and the Cooking activity.

Once the prediction has been made, the DMS will use these rankings (and the associated score to each label in the ranking) in two different ways.

- On the one hand, if the robot behaves proactively, it will make three hierarchical decisions: category, subcategory, and activity that will finally execute. Figure 9.6 shows an example (in highlighted in red) of a DMS selection. In this case, the robot selects the category *Multimedia*, the subcategory *Videos*, and finally launches the activity that displays a video showing a cooking recipe.

- On the other hand, if the robot cedes the initiative to the user allowing them to select the entertainment activity, the robot will show a sequence of tablet menus where the preferred items of each category and subcategory will appear on top of non-preferred ones. This organisation will facilitate the users' selection of their preferred activities.

It is worthy to note that the selection process does not always choose the items of each ranking with the higher score. Instead, and using the Boltzmann's action selection method described in Appendix B.1, random selection combines with the selection of preferred activities to promote exploration. In our model, all items (and activities) have a rating value from 0 to 5, where higher ratings mean the user prefers that item above low rated ones.

9.2.2. Experimental setup

Since LR is a machine learning method used for ranking generation based on Random Forest, the model built for this application requires training data for estimating the user preferences. Thus, the training data consisted of an input vector of user features with labelled output data containing the preferences towards the entertaining activities of Mini. As in most classification problems, the biggest the dataset, the better the estimations produced by the model. Due to the Covid-19 pandemic, data retrieval could not be done face to face. Instead, we opted for using an online survey for building up our dataset. Using similar studies in data retrieval [340, 341], the survey contained questions about three different topics: personal data, habits/interests, and preferences towards the entertainment activities of the robot. As Figure 9.7 shows, the users' personal questions comprised their age, sex, education level, and similar attributes. The answers to these questions were, in all cases, limited to a set of predefined options. The answers to the questions about habits/interests were limited to yes/no answers, and questions related to preferences were limited to a 5-point Likert scale where 1 meant do not like the activity and 5 the user likes the activity a lot.

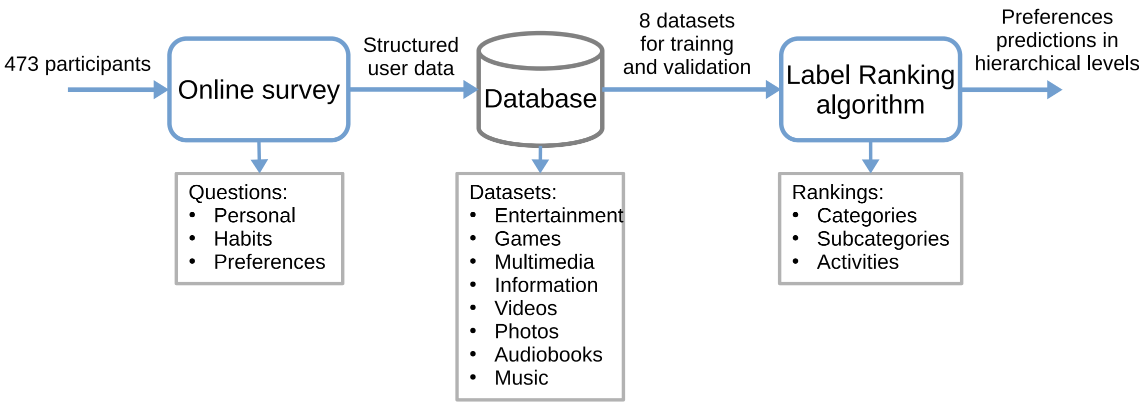


Figure 9.7: Data collection and use in the preferences estimation process.

473 participants (241 men, 217 women aged $\mu = 29.52, \sigma = 12.64$) completed the online survey. We had to remove the answers of two participants due to the inconsistencies found in their answers (for example, one user reported 6 years old and a bachelor's degree). Valid answers defined the input space of the model, being used for training and evaluating its accuracy in predicting rankings. In total, using the replies provided by the survey participants, eight datasets were defined organised in three hierarchical levels.

As Figure 9.6 shows, our learning problem contains 8 datasets (entertainment, games, multimedia, information, videos, photos, audiobooks, and music), organising the categories, subcategories, and activities in hierarchical order. On top of the hierarchy, the first dataset ranks the categories games, multimedia, and information. In the middle of the hierarchy, the second dataset ranks videos, photos, audiobooks, music, jokes, and sayings classified in the multimedia category. Finally, in the lower level, a dataset ranks different kinds of games, another dataset sorts the information applications, and four additional datasets rank the activities that belong to the multimedia subcategories: videos, photos, audiobooks, and music. As mentioned before in this dissertation, this organisation allows the decision-maker to select the category, subcategory, and activity that the robot will finally execute using a hierarchical method based on Boltzmann's equation (see Appendix B). This action selection method allows us to foster action exploration and vary the action selection process by intercalating proactive autonomous decisions of the robot with partial user selections. The Boltzmann equation assigns a selection probability to each action depending on a score representing the optimal action and a parameter that regulates the randomness in the selection.

Due to the wide range of LR algorithms in the literature, we opted for comparing two promising ones to obtain which one produces the best results in terms of ranking accuracy and computational payload. We selected Top Label as Class (TLAC) and Ranking by PairWise Comparison (RPC) (see Section 4.3.2.1 for further details) because both provide excellent results in ranking preferences, they use different approaches to produce their predictions, and their computational requirements are below the other alternatives. Due to this work's particular application (social robotics), we wanted to find which one suits better in our scenario. In the comparison process, we included hyperparameter optimisation mechanisms to attain the best performance of both algorithms.

After finding the best LR algorithm for our application, we performed a HRI study to investigate whether people really perceive the robot's personalisation during the performance of entertainment activities. In the study, 22 participants (13 women, 9 men aged $\mu = 41.09, \sigma = 20.57$) were equally divided in two conditions:

- Condition 1, *personalised robot*: the robot estimated three activities using the preferences predictions during a short-term interaction using our proposed LR method.
- Condition 2, *general robot*: Mini selected three random activities during a short-term interaction with the user, not including any personalisation.

After completing the activities, the participants had to complete a survey to evaluate whether the *personalised robot* (Condition 1) was better considered in terms of adaptation, appropriateness, and intelligence than the *general robot* (Condition 2). The following section details the evaluation of the experiment we conducted to validate the personalisation of entertainment activities proposed in this work.

9.2.3. Evaluation

The experiments presented in the previous section were evaluated using well-known questionnaires and statistical analysis.

- On the one hand, the comparison of the LR algorithms was carried out in terms of the prediction accuracy, the execution time for running the training of the learning model, and ranking correlation metrics. In this work, we used two well-known ranking correlation metrics Kendall's τ-b [342], and Spearman's ρ [343] (see Appendix A for further information of both metrics) for measuring the correlation between the real rankings and the estimated ones. Using a second metric (Spearman's ρ) supports the results obtained with the first metric (Kendall's τ-b). Kendall τ-b is more meticulous in comparing the rankings, usually yielding lower correlation values than Spearman ρ. These metrics provide a value in the range $[-1, 1]$ that indicates the similarity between the two rankings. According to previous studies in statistics [344, 345], a positive rank correlation below +0.3 is considered weak, from +0.3 to +0.7 is considered moderate, and above +0.7 is considered strong. A negative correlation value indicates that the predicted ranking is reversed for the real ranking. The validation of both TLAC and RPC was produced using 10-fold cross-validation [346]. The tuning of the hyperparameters of the model used a grid search method [347] (see Appendix B for a brief description of both methods).

- On the other hand, the evaluation of the second experiment consisted of using the widespread Godspeed questionnaire [348] (included in Appendix D.5) to obtain the participants' perception of the robot. We designed two scenarios:

  - Interacting with a personalised robot suggests predicted activities (Condition 1).
  - Interacting with a general robot without personalisation (Condition 2).

  Since the evaluation was conducted with Spanish participants, we used the official translation of the questionnaire[4] [349]. The questionnaire obtains the user's impressions about the robot in five areas: Anthropomorphism (A), Animacy (A), Likeability (L), Intelligence perceived (IP), and Security perceived (SP).

  We complemented the Godspeed questionnaire with six ad-hoc questions (see Appendix D.2 for their Spanish translation) about the personalisation and adaptation process using previous similar works as models [350, 351]. These questions were aggregated into the attribute *Personalisation perceived*, which measured the participant's perception about the robot's personalisation to their features. The user answered these questions using a 5-point Likert scale where 1 means strongly disagree and 5 strongly agree. The ad-hoc questions are listed below.

  - **Q1:** How appropriate did you find the first activity selected by the robot according to your preferences?

[4]Godspeed questionnaire official website.

- **Q2:** How appropriate did you find the second activity selected by the robot according to your preferences?
- **Q3:** How appropriate did you find the third activity selected by the robot according to your preferences?
- **Q4:** In general, do you think the activities proposed by the robot are adequate to you?
- **Q5:** Have you noticed that the robot knows information about yourself?
- **Q6:** Have you noticed that the robot behaviour adapts to yourself?

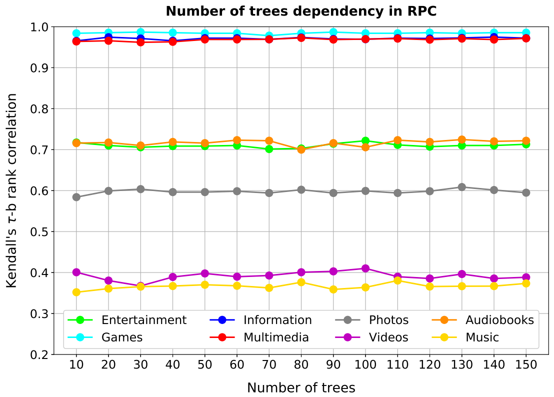
According to our hypothesis, those participants using the personalised robot should perceive it as more appropriate. Considering the categories assessed by the Godspeed questionnaire [348], we expect to obtain significant statistical differences in the Animacy (A), Likeability (L), and Intelligence perceived (IP). Besides, we expect to find significant statistical differences in the Personalisation perceived attribute, indicating that users using the robot with personalised activity selection perceive the robot as more appropriate to their likes.

9.2.4. Results

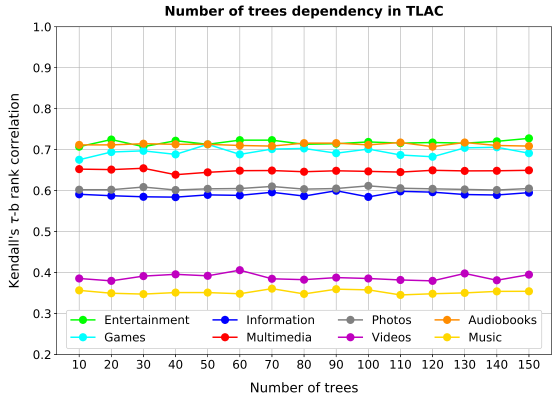
Following, we present the results of the two experiments we conducted for evaluating the estimation of user preferences for HRIs. First, we show the comparison of TLAC and RPC, reaching a conclusion about which algorithms works better in our application. Second, we show the results of the HRI study for assessing the activity selection personalisation.

9.2.4.1 Comparison results

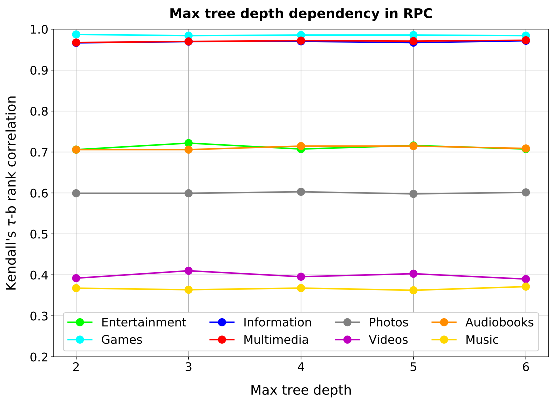
The first step in obtaining the best performance of each LR algorithm was to find the best combination of hyperparameters. Figures 9.8 show the hyperparameter optimization of TLAC and RPC in terms of the number of trees of the forest n_{trees} and the maximum depth of the forest m_{depth}, maintaining a fixed value for the hyperparameters minimum samples in a non-leaf node $n_{samples}$, and the minimum samples in a leaf node n_{leaf}. The optimisation was carried out using the grid search method [347]. Besides, we studied the impact of splitting the nodes of the random forest using the Gini impurity (see Appendix A) or information gain (see Appendix B) methods. Like these results show, RPC produces the best results using $n_{samples} = 10$, $n_{leaf} = 5$, $m_{depth} = 3$, and $n_{trees} = 100$. On the other hand, TLAC produces the best outcomes using $n_{samples} = 15$, $n_{leaf} = 10$, $m_{depth} = 4$, and $n_{trees} = 60$. The hyperparameters that produce the best results for each algorithm are used to attain the best performance of each dataset in the ranking process. Note that we obtained these optimal hyperparameters considering the eight datasets that define the preferences rankings of our system grouped in categories, subcategories, and groups of activities. Since looking for the specific optimal parameters of each dataset is intractable due to the large dimension of the system, we opted for concentrating on getting those hyperparameters that provide the best results for all the datasets on average.



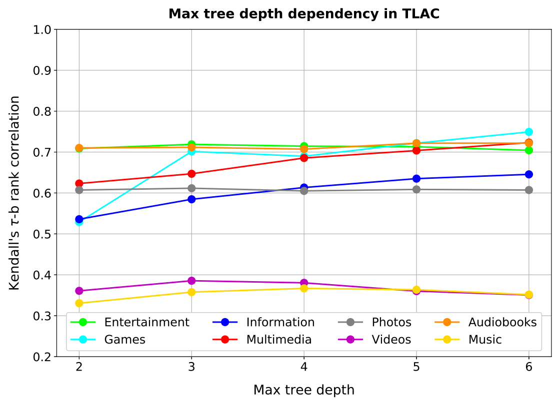
(a) Kendall τ-b average value depending on n_{trees} of the forest in RPC algorithm with fixed values of m_{depth} = 3, $n_{samples}$ = 15 and n_{leaf} = 10.



(b) Kendall τ-b average value depending on n_{trees} of the forest for TLAC algorithm with fixed values of m_{depth} = 4, $n_{samples}$ = 10 and n_{leaf} = 5.



(c) Kendall τ-b average value depending on m_{depth} of the tree in RPC algorithm with fixed values of n_{trees} = 100, $n_{samples}$ = 15 and n_{leaf} = 10.



(d) Kendall τ-b average value depending on m_{depth} of the tree in TLAC algorithm with fixed values of n_{trees} = 60, $n_{samples}$ = 10 and n_{leaf} = 5.

Figure 9.8: Representation of n_{trees} and m_{depth} hyperparameters optimization on each dataset considering the set of hyperparameter that provided the best results in each trial.

Once we found the optimal hyperparameter for the model, we validated the results in terms of the classification accuracy of the forest used in the classification process, the average execution time required for training the model using 10-fold cross-validation, and the ranking correlation metrics Kendall τ and Spearman ρ considering each algorithm and each dataset. Table 9.1a shows the results concerning the classification accuracy, Table 9.1b shows the training average time required by each dataset, Table 9.1c shows the Kendall τ rank correlation metrics for each dataset, and Table 9.1d shows the results of the Spearman ρ rank correlation metric. These tables show the average values (an standard deviation in brackets) obtained from the 10-fold cross-validation of each algorithm considering each dataset.

Dataset	RPC Gini Impurity	RPC Information Gain	TLAC Gini Impurity	TLAC Information Gain
Entertainment	**.8630 (.0227)**	.8630 (.0308)	.7326 (.0583)	.7347 (.0788)
Games	.9934 (.0068)	**.9942 (.0054)**	.9804 (.0180)	.9934 (.0099)
Multimedia	.9862 (.0047)	**.9872 (.0044)**	.9217 (.0325)	.9500 (.0377)
Information	**.9886 (.0053)**	.9881 (.0044)	.9304 (.0474)	.9000 (.0468)
Photos	**.8057 (.0137)**	.8050 (.0283)	.7456 (.0937)	.7500 (.0668)
Videos	**.7065 (.0191)**	.7050 (.0222)	.6239 (.0558)	.6282 (.0655)
Audiobooks	.8666 (.0269)	.8681 (.0252)	.8760 (.0292)	**.8826 (.0403)**
Music	**.6962 (.0226)**	.6928 (.0154)	.3847 (.0550)	.3739 (.0946)

(a) Mean (std) average values for internal Random Forest classifier accuracy obtained using optimal hyperparameters in 10-fold cross validation.

Dataset	N Labels	RPC Gini Impurity	RPC Information Gain	TLAC Gini Impurity	TLAC Information Gain
Entertainment	3	3.8511	3.8414	2.4019	**2.2970**
Games	3	4.1858	4.0515	**2.0990**	2.1385
Multimedia	6	18.3387	17.8568	**3.3404**	3.3899
Information	6	18.5684	18.4749	4.4734	**3.8260**
Photos	4	7.2872	7.2957	2.8319	**2.6490**
Videos	4	7.9843	8.0638	2.5251	**2.4801**
Audiobooks	3	3.7682	3.9130	2.3731	**2.3180**
Music	7	26.7077	26.6414	5.5840	**5.0993**

(b) Mean execution time values per each cross-validation iteration (training and validation), which depend on the number of labels of each ranking using optimal hyperparameters.

Dataset	RPC Gini Impurity	RPC Information Gain	TLAC Gini Impurity	TLAC Information Gain
Entertainment	**.7260 (.0455)**	.7260 (.0616)	.7246 (.0349)	.7246 (.0685)
Games	.9869 (.0136)	**.9884 (.01084)**	.7637 (.0954)	.8362 (.0534)
Multimedia	.9721 (.0113)	**.9753 (.0127)**	.7315 (.0286)	.7292 (.0223)
Information	**.9762 (.0112)**	.9744 (.0090)	.6530 (.0286)	.6997 (.0314)
Photos	**.6115 (.0275)**	.6115 (.0584)	.6188 (.0897)	.6137 (.0471)
Videos	.4007 (.0412)	**.4094 (.0538)**	.4057 (.0376)	.4014 (.0669)
Audiobooks	.7333 (.0539)	**.7362 (.0505)**	.7289 (.0439)	.7318 (.0481)
Music	**.3805 (.0460)**	.3799 (.0284)	.3726 (.0427)	.3788 (.0338)

(c) Mean (std) Kendall's τ-b average values for 10-fold cross-validation obtained for optimal hyperparameters.

Dataset	RPC Gini Impurity	RPC Information Gain	TLAC Gini Impurity	TLAC Information Gain
Entertainment	.7576 (.0462)	**.7608 (.0631)**	.7554 (.0315)	.7521 (.0660)
Games	.9902 (.0102)	**.9913 (.0081)**	.8206 (.0711)	.8739 (.0438)
Multimedia	.9831 (.0069)	**.9849 (.0090)**	.8108 (.0290)	.8126 (.0221)
Information	**.9853 (.0081)**	.9848 (.0062)	.7636 (.0248)	.7983 (.0286)
Photos	.6765 (.0285)	.6765 (.0641)	**.6782 (.0881)**	.6734 (.0468)
Videos	.4465 (.0555)	.4486 (.0590)	**.4630 (.0328)**	.4578 (.0681)
Audiobooks	**.7717 (.0500)**	.7706 (.0450)	.7684 (.0434)	.7706 (.0436)
Music	.4781 (.0550)	.4754 (.0330)	.4760 (.0563)	**.4835 (.0333)**

(d) Mean (std) Spearmans's ρ average values for 10-fold cross-validation using optimal hyperparameters.

Table 9.1: Results in terms of the accuracy of the random forest classifier, average execution time, Kendall's τ-b, and Spearman's ρ ranking correlation metrics. These metrics are presented for Top Label as Class (TLAC) and Ranking by PairWise Comparison (RPC) algorithms considering the gini impurity and the information gain as splitting criteria. **Bolded numbers** indicate the best performance for each dataset, algorithm, and splitting criteria.

Considering the Table 9.1a that represents the classification accuracy of the random forest for both LR algorithms, it is possible to observe clear tendencies in the data. First, RPC outperforms TLAC in all datasets excepting in the audiobooks one. Second, it is possible to notice that Gini impurity produces better node splitting than Information gain in most cases, and in those where Information gains work better, differences are minimal. Finally, the classification accuracy is quite good for all datasets, presenting values above 70% of accuracy. Besides, the classification accuracy in some datasets (games, multimedia, and information) is excellent, above 98%.

Table 9.1b shows the average execution time (in seconds) for each algorithm, splitting criteria, and dataset considering the number of labels to rank. As expected, TLAC provides faster predictions than RPC. Since RPC trains $N - 1$ models instead of 1 (TLAC), the number of labels in the ranking substantially slows the training of the system. However, the maximum time taken is around 260 seconds for the 10 cross-validation trials, a time that can be easily afforded if the training occurs during the night or when the robot is relaxed.

Analysing the results of the ranking correlation metrics represented in Tables 9.1c and 9.1d, it is possible to make important conclusions. Like Table 9.1c shows, Kendall $\tau - b$ values are, on average, substantially better for RPC than for TLAC in three of the datasets (games, multimedia, and information). Besides, in our application, RPC performs better in the rest of the datasets than TLAC. According to the criteria used for splitting the trees, no significant differences between Gini impurity and Information gain.

Focusing on the best values for each dataset, the evaluation results look promising since all datasets present a moderate rank correlation. Outstanding results are obtained in the games (+0.9884), multimedia (+0.9753), information (+0.9744), entertainment (+0.7260), and audiobooks (+0.7362), where a strong positive rank correlation (above +0.7) is obtained when comparing predicted and real rankings. In the other datasets, a moderate rank correlation (from +0.3 to +0.7) is obtained (photos (+0.6115), videos (+0.4094), and music (+0.3805)).

Comparing the classification accuracy results represented in Table 9.1a and ranking correlation metrics shown in Table 9.1c and 9.1d it is possible to perceive that the accuracy of the classification stage influences the rank aggregation methods carried out for obtaining the final predictions of the algorithms. This fact means that a correct classification is critical in the success of the ranking prediction, is also affected by each dataset's distribution and size.

The Spearman ρ ranking correlation metrics provided very similar results to the Kendall τ-b (see Table 9.1d). Nevertheless, the values provided by Spearman ρ are slightly above Kendall τ-b ones because Kendall τ-b is more meticulous in comparing the estimated and real rankings. For this reason, although TLAC gives better results than in Kendall τ-b metric, RPC is still the algorithm with the best performance. The ranking correlation in terms of the Spearman ρ metric supports the previous results since most datasets have a positive correlation above 0.4, considered moderate, reaching values above +0.9 in some cases.

Considering the previous results, we can conclude that RPC is the best alternative for our application. On average, the outcomes provided by RPC are subtly above the ones provided by TLAC either on the prediction accuracy and ranking correlation metrics. For this reason, we opted for using RPC in our application and the HRI study presented in the following section.

9.2.4.2 HRI study

The HRI study to appraise whether the personalisation of the interaction with the users affects the perception and acceptance of the robot provided positive results. As Table 9.2 and Figure 9.9 show, significant statistical differences appeared when comparing the robot personalising the activity selection (Condition 1) and the robot making random activity selection (Condition 2) in two of the five categories of the Godspeed questionnaire. The analysis was carried out using the well-known T of Student test for data following a normal distribution with a small sample size. The study yielded significant statistical differences in the Godspeed attributes Likeability (L) (p-value < 0.001) and Intelligence perceived (IP) (p-value < 0.001). Nevertheless, contrary to our expectations, the Animacy (A) category did not differ between both conditions. Neither the categories Anthropomorphism (A) nor Security perceived (SP).

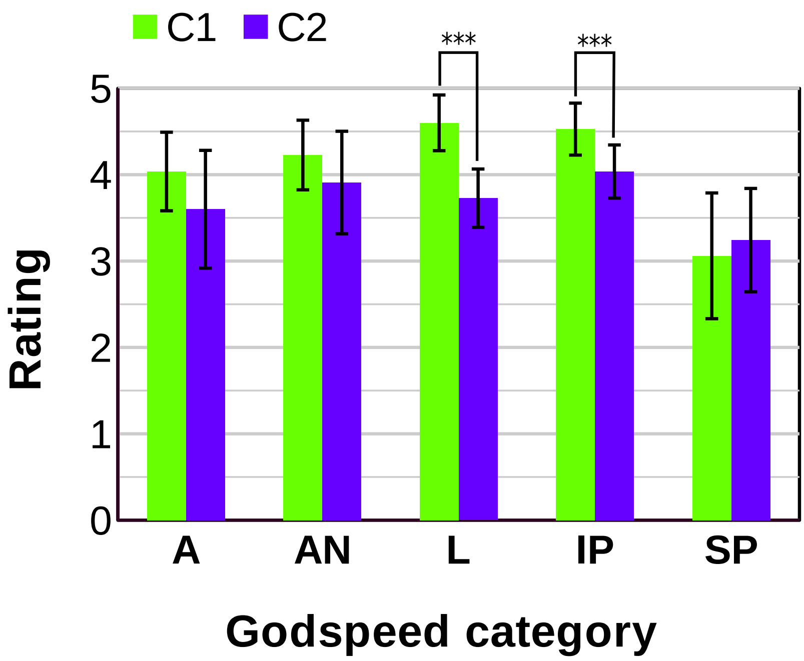


Figure 9.9: Mean and standard deviation values in the statistical analysis of the Godspeed results considering each category and condition tested Anthropomorphism (A), Animacy (A), Likeability (L), Intelligence perceived (IP), and Security perceived (SP). Significant statistical differences (p-value from 0.05 to 0.01) are denoted with $*$, very significant (p-value from 0.01 to 0.001) with $**$, and extremely significant (p-value$<$ 0.001) with $***$.

Goodspeed category	Students' T	DoF	p-value
Anthropomorphism	1.767	5	.092
Animacy	1.471	6	.157
Likeability	6.197	5	.000***
Intelligence perceived	3.788	5	.000***
Security perceived	0.640	3	.529

Table 9.2: Statistical results of the five Godspeed categories comparing the two conditions assessed. Significant statistical differences (p-value from 0.05 to 0.01) are denoted with $*$, very significant (p-value from 0.01 to 0.001) with $**$, and extremely significant (p-value$<$ 0.001) with $***$.

Regarding the *Personalisation perceived* attribute oriented to measure whether users perceive the selection of the robot appropriate to them and whether they perceived the adaptive robot interaction, Table 9.3 and Figure 9.10 show the results obtained in the statistical analysis of both conditions. The outcomes provided by the Mann-Whitney's U [352] statistic for non-normal distributions show significant differences between both conditions (p-value = 0.003). These results suggest that the study participants perceived the robot personalisation and, based on the Godspeed results, they perceived it as more intelligent and likeable.

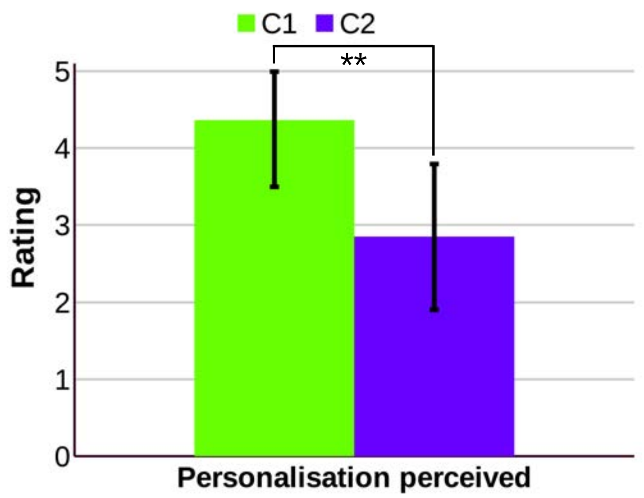


Figure 9.10: Mean and standard deviation values of Personalisation perceived attribute for each condition tested. Significant statistical differences (p-value from 0.05 to 0.01) are denoted with $*$, very significant (p-value from 0.01 to 0.001) with $**$, and extremely significant (p-value$<$ 0.001) with $***$.

Attribute	Mann-Whitney's U	DoF	p-value
Personalisation perceived	16.50	5	.003**

Table 9.3: Statistical results using the Mann-Whitney's U score for the Personalisation perceived attribute. Significant statistical differences (p-value from 0.05 to 0.01) are denoted with $*$, very significant (p-value from 0.01 to 0.001) with $**$, and extremely significant (p-value$<$ 0.001) with $* * *$.

9.2.5. Discussion

The results of this second experiment showed that despite TLAC requires less computational time for training and validating the model, and its performance is worse than the outcomes produced by RPC. Statistically, the ranking correlation metrics show promising results in the prediction of the rankings, presenting in all the cases positive moderate and robust correlations during the validation process. However, the exponential increase in the training time and the decrease of the ranking correlation values required by the datasets with a large number of labels to rank (e.g. music with 7 labels) suggest that LR algorithms are highly dependent on this factor, decreasing their efficiency as the number of labels to rank increases.

The two algorithms compared in this experiment provide outstanding results in the ranking estimation. However, we opted for selecting RPC above TLAC with Gini impurity as a splitting method since its the alternative that generally produces the best results in terms of prediction accuracy and the ranking correlation metrics. However, this alternative requires more computational time due to the number of models built for training the system. Despite the training time exponentially increasing with the number of labels to rank, this application is not constrained by time limits, as training can be produced during the low demanding of the robot.

The contribution of the PL System to social robotics lies in the use of LR methods to endow these systems with the capability to personalise the interaction with humans by suggesting them their favourite activities instead of selecting them randomly. This approach, previously tackled in recommender systems but not in social robots, must be tested in real human-robot scenarios. For this reason, we conducted a second experiment to assess whether the robot users perceived the robot that could personalise its entertaining activities selection more adaptive to them. The results of this HRI experiment showed that participants using a personalised robot could perceive such personalisation and also perceive the robot as more likeable and intelligent.

9.3. Adjusting the initial estimations

The third step of the PL System comprised the adjustment of the initial estimations using RL and the feedback provided by the user while performing the activities. Next, we describe this second approach, the experiment we designed, its evaluation, and results.

9.3.1. Defining the learning scenario

Although estimating the user's preferences gives the robot an initial approximation to suggest their potentially preferred activities, these predictions might not be their actual preferences. For this reason, seeking to improve the initial learning system, we extended the PL environment by including a RL method that allows the robot to tune up these initial estimations using the interaction with the user during the robot's typical operation. This third step uses the user's feedback to adjust the score associated with each ranking label. Mini obtains the user's feedback using two different but complementary mechanisms.

- Directly asking the users how much they liked a specific activity after executing it (direct feedback).
- Estimating the feedback using interaction metrics representing the quality of the interaction and user engagement (indirect feedback).

The execution of the entertainment activities and the user's feedback helps attain a more representative characterisation of the user's preferences. Besides, it is worthy of mentioning that in most RL applications, the learning system is expected to attain an optimal solution. However, and considering that the user's preferences may vary in the long run, in this scenario, we do not expect to attain convergence but to continuously adapt the scores of each label to perceive the changes in the user's likings.

We evaluated this methodology in two independent experiments where we appraised the processes of estimating the user's preferences and adjusting the predicted scores. Next, we present both experiments giving critical remarks about how they were performed.

9.3.2. Experimental setup

The third set of experiments had the goal of researching whether the adaptation of the initial predictions using RL produce a good representation of the authentic preference of each user. Besides, we compared the performance of the adaptive preference system using three different approaches:

- Using direct feedback: the robot obtains direct feedback by asking the user how much they liked the activity after executing them. Using a 5-point Likert scale [353, 354] the user rates their preferences to the activities of the robot.

- Using indirect feedback: the robot estimates indirect feedback using different metrics retrieved from the interaction flow.

- Combining direct and indirect feedback.

We considered the entertainment activities related to music, photos, and videos subcategories for this experiment. Figure 9.11 shows the entertainment activities used in this experiment.

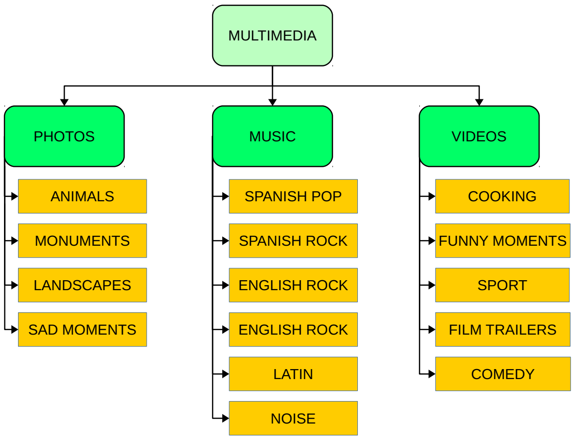


Figure 9.11: Set of activities used in the evaluation of the adjustment phase in the Preference Learning System.

Eight participants (2 female, 6 male) were divided into the three conditions. Three participants realised the condition of direct feedback, another three realised the condition of indirect feedback, and the two remaining participants did the mixed condition (combination of direct and indirect feedback). Before starting the experiment, we requested the participants to fill out a survey indicating how much they liked each activity evaluated in the experiment. This survey took place before the participants tested the robot's activities. This survey provided us with the real preferences of each participant, allowing us to compare them with the results attained by the adaptive RL system. Then, the experiment consisted of the participants interacting with the robot in sessions of 20 minutes. The robot randomly decided the entertaining activities the user had to perform, promoting a balance in the activity selection. The idea was to explore as many activities as possible to learn the corresponding score of each activity. While performing an activity, the user could cancel it by touching the robot's right shoulder. The experiment finished when all participants completed at least 14 sessions with the robot within a month and a half. Some participants completed more sessions than 14.

It is worthy to stand out as an essential aspect of the experiment that in the condition where only direct feedback was used, the question to obtain the user's rating was not always presented after an entertainment activity to avoid fatiguing the user. Instead, the question was presented in half of the activities performed. If the question was not presented, the activity's associated score (Q(a) or Q-value) was not updated. It is also worth remembering that the updating of the Q-value of each activity was carried out using Equation 9.1, a formula derived from the action-based adaptive RL previously introduced in Section 4.3.2.2.

$$Q(a) \leftarrow Q(a) + \alpha \left[r - Q(a) \right] \tag{9.1}$$

As the previous equation defines, the Q-value (Q(a)) update requires a reward value (r) obtained from the interaction that indicates how much the user liked the activity. Equation 9.2 defines the direct feedback reward value (r_{direct}) of this problem.

$$r_{direct} = \text{User rating} \in \{1, 2, 3, 4, 5\} \tag{9.2}$$

The user's question to rate the activity never appeared in the condition using only indirect feedback. Instead, we used the interaction metrics to estimate the reward value. In this condition, the score (Q-value) was continuously updated after executing the activity since the interaction metrics do not require the user's intervention. In this experiment, we opted for using the following interaction metrics:

- User selection (US): This metric represents if the user selected the activity (1) or if it was autonomously proposed by the robot (0).
- Activity Result (AR): All activities can have two possible outcomes: succeeded, in whose case the value of this metric is 1, or aborted, an event that is provoked by the user when (s)he voluntary stops the activity, in which case the value of the metric is 0.
- Execution time (ET): This metric represents if the execution time of the activity is inside a specific range concerning the average execution time of the rest of the users. If the execution time of the activity is inside the range $[\mu - \sigma, \mu + \sigma]$ (where μ is the mean execution time of the activity considering all the users and σ the standard deviation), the value of the metric is 1. Otherwise, the value is 0.

Each metric presented above can have two possible values: 1 if the condition indicated by the metric is true, or 0 otherwise. As Equation 9.3 shows, the value of each metric defines a reward value ($r_{indirect}$) from 0 to 5 units used for updating the Q-value of the activity just performed. Thus, Equation 9.3 represents the indirect feedback reward function of this RL problem.

$$r_{indirect} = 2 \cdot US + 2 \cdot AR + ET \tag{9.3}$$

Finally, in the condition combining both types of feedback, the question for obtaining the direct feedback was not always presented to the user but had the same probability of being or not being asked after the execution of each activity. Nevertheless, indirect feedback was always used since the interaction metrics always produce estimations. At the beginning of the experiments, the scores of all the activities valued 5 units to homogenise the adaptation process. Equation 9.4 presents how the reward value (r) was calculated for each of the conditions evaluated in the experiment.

$$r = \begin{cases} r_{direct} = \text{User rating} \in \{1, 2, 3, 4, 5\} & \text{Direct feedback} \\ r_{indirect} = 2 \cdot US + 2 \cdot AR + ET & \text{Indirect feedback} \\ (r_{direct} + r_{indirect})/2 & \text{Combined feedback} \end{cases} \tag{9.4}$$

We hypothesised that the condition combining both types of feedback will produce the best results since it does not saturate the user with endless questions about rating the activities and because the combination of both feedbacks speeds up the adaptive process. It is important to note that in those setting where both types of rewards are combined, the reward function averages the value obtained Equations 9.2 and 9.3, as Equation 9.4 indicates. The following Figure 9.12 shows the dynamics of the experiment carried out for adapting the preference of the user considering the condition combining direct and indirect feedback.

9.3.3. Evaluation

The evaluation of the preferences adjustment performed from the estimation made by RPC using direct and/or indirect feedback was carried out employing the following metrics (Appendix A provided further details of these metrics).

Departing from the previous experiments in which we evaluated the positive effects of personalising HRI, we evaluated whether the adaptation of the initial predictions improved the system by perceiving possible changes in the users' likings. To facilitate the comprehension of the results, we first show representative cases yielded by adapting each activity's scores (Q-values). Then, we compare the differences between the three conditions used for obtaining the user's feedback needed to adapt the scores: using direct feedback asking the user how much (s)he liked the activity, estimating the feedback using interaction metrics, or the combination of both. We use two metrics for this evaluation:

- First, we measure the Relative Standard Deviation (RSD) to obtain the differences between the final score learnt by the system to the real values provided by the user.
- Then, the Root Mean Squared Error (RMSE) provides a most acceptable representation of the error committed by each condition in the adaptation of the scores regarding each activity.

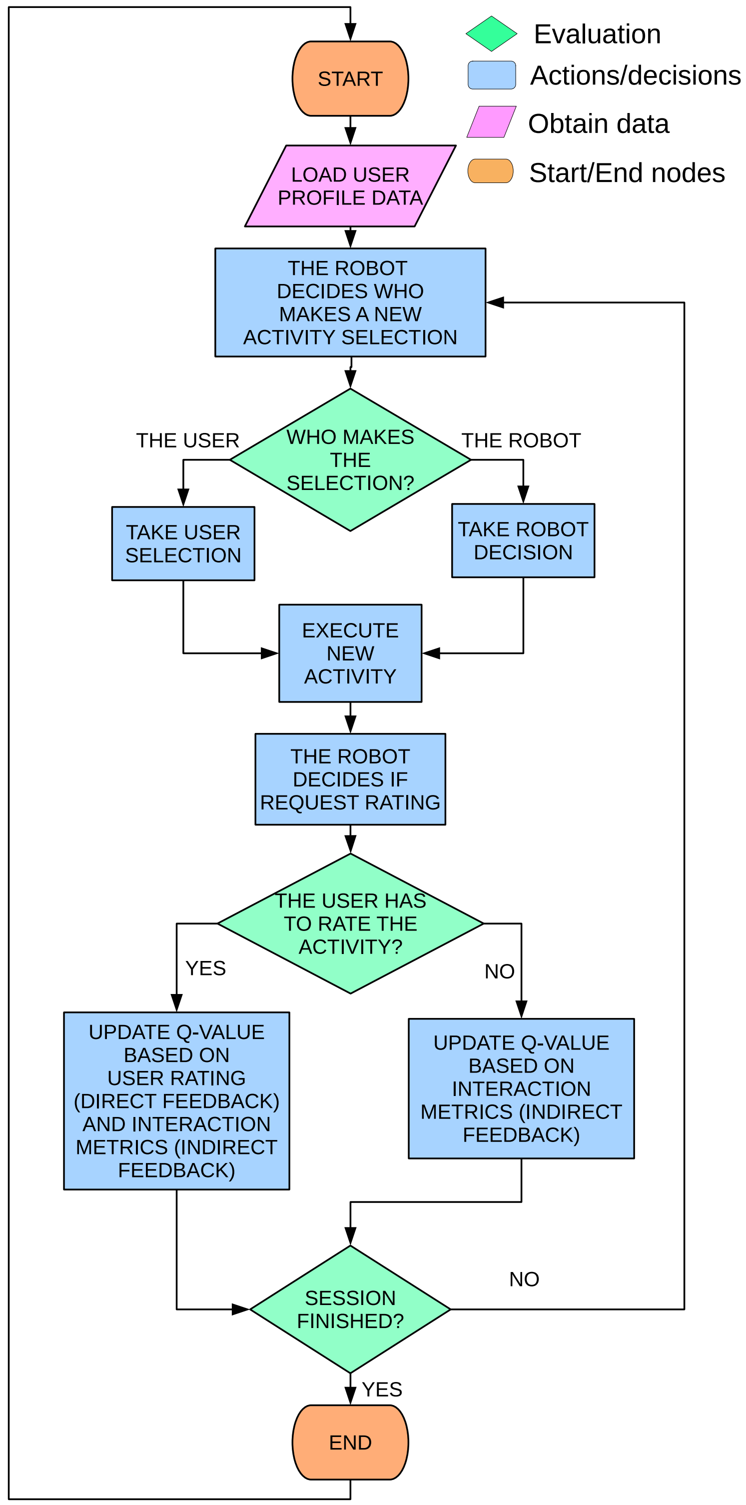


Figure 9.12: Diagram showing the dynamics of the user preferences adaptation experiment.

9.3.4. Results

Following, we present the results obtained during the learning and evaluation of the preferences adjustment using RL.

9.3.4.1 Representative cases

Next, we describe the most representative cases attained during the adjustment of the Q-values using RL.

Perfect fitting:

The first representative case shows, in Figure 9.13, the perfect adaptation of a Q-value. In this case, the participant was classified in the condition using direct feedback. The user consistently rated the activity with the same grade (1) during several interactions, making the Q-value converge to a final value. As Figure 9.13 shows, the final Q-value reached (blue) coincides with the initial value indicated by the participant in the preliminary survey (red). This case corresponds to the activity showing photos of sad moments, so it seems coherent that the user did not like it. Although the evolution of the Q-value smoothly converges to a final value of 1 units, it is important to say that if the user changes his opinion about the activity in future interactions, the new Q-value will move towards the new rating provided by the user.

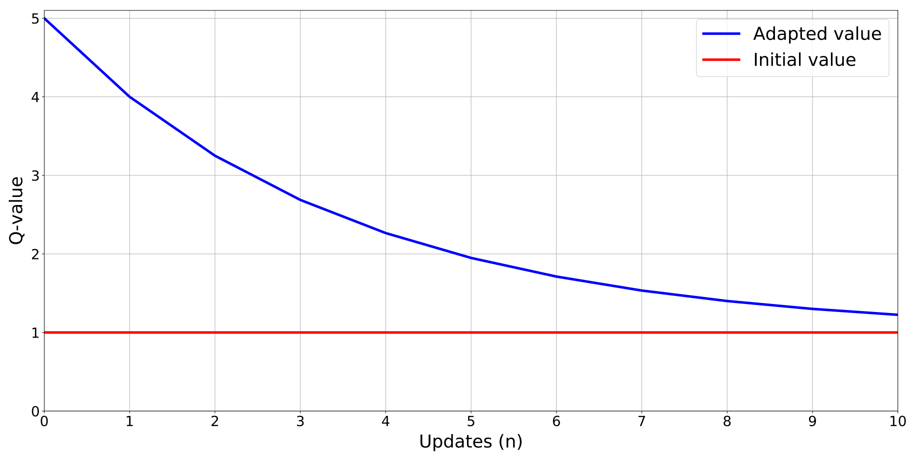


Figure 9.13: Smooth and perfect adaptation to the initial preference of a user classified in the condition where just direct feedback was considered.

Irregular fitting to a new value:

The second representative case shows in Figure 9.14 an adaptive process where the final Q-value reached by the system (blue) does not coincide with the initial indication of the user in the preliminary survey (red). Besides, as the adaptive processes combine direct and indirect user feedback, the evolution of the Q-value is more irregular than in the previous example. Although the irregularity of the adaptation, it is important to note the higher number of updates of this case about the previous one, since the combination of two types of feedback allows to retrieve more information of the user leading to a faster adaptive process.

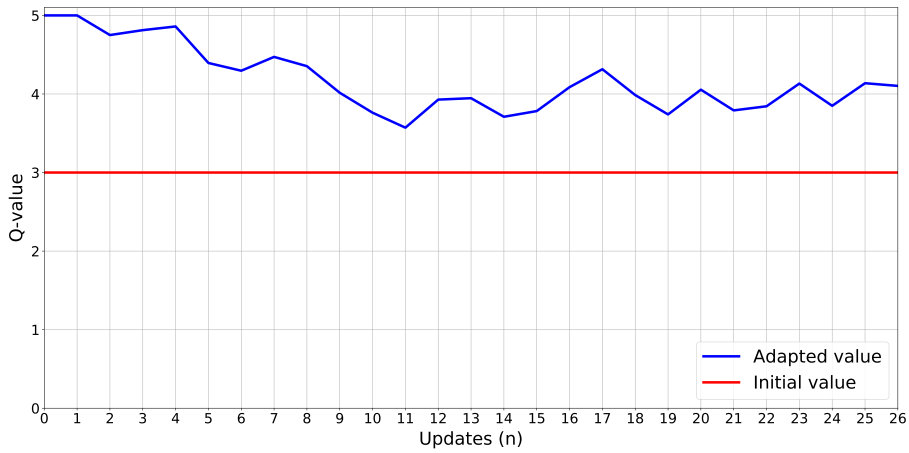


Figure 9.14: Irregular but faster adaptive process to a value that is not the initially indicated by the user. This features of this evolution respond to the combination of direct and indirect feedback.

Irregular fitting to initial value:

The third representative case, depicted in Figure 9.15, shows a Q-value evolution characterised by an irregular but correct adaptation. The evolution finally attains the Q-value initially indicated by the user in the preliminary survey (red) although it dithers a little bit. The indirect feedback provokes that the estimated Q-values are not regular since the user's interaction metrics are estimated and not provided, leading to a weak variation in the predicted rewards using the interaction metrics.

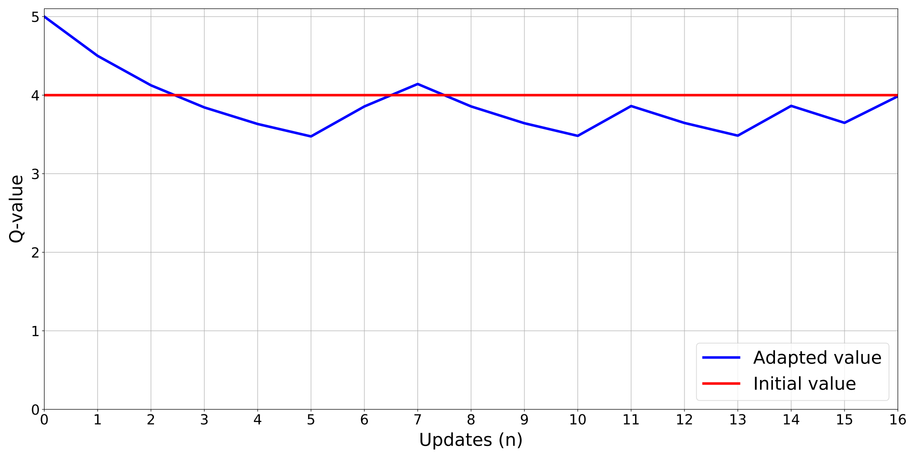


Figure 9.15: Adaptive evolution of the Q-value just using indirect user feedback.

Not convergent fitting:

Finally, the fourth representative case shows in Figure 9.16 an adaptive process that does not converge to any final Q-value (blue). However, the evolution shows how the Q-value gets closer to the initial value indicated by the user (red), the irregular ratings of the activity provided by the user and the estimated rewards using the interaction metrics. Although it may seem that this adaptive case is incorrect, it is due to a change in the user's preferences during the realisation of the experiment since the rewards provided were very different from each other. This example represents exactly what we pretend to reach using a continuous adaptive system, that in case the user changes its opinion about an activity, the robot can recognise it and adapt to it.

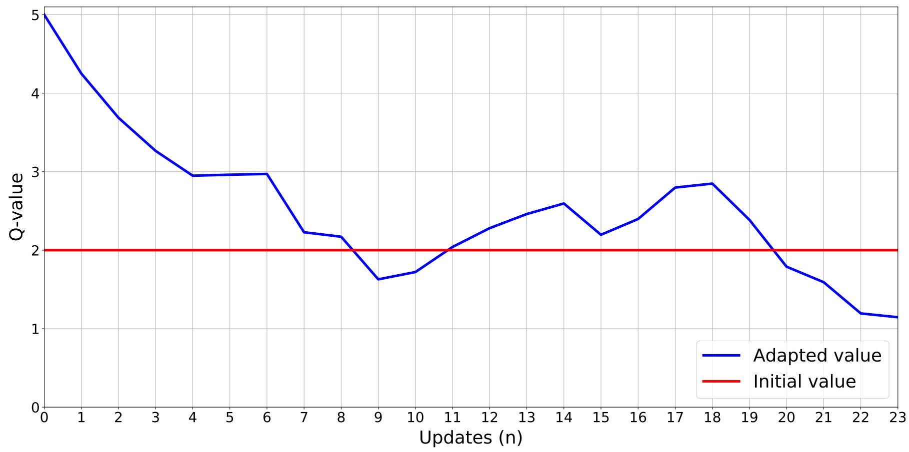


Figure 9.16: Very irregular evolution of a Q-value that does not converges to any specific value at the end of the trial.

9.3.4.2 Comparing the evaluation metrics

For the understanding of the three conditions evaluated in the user preferences adaptive process, Table 9.4 is fundamental. This table contains the number of updates of each action, the final Q-value of each activity, and the initial score provided by the participants before the experiment. Besides, it includes the number of sessions of 20 minutes of duration done by each participant. Highlighted in blue, those Q-values whose number of updates are below 3 updates. In red the four representative case described in the previous section.

Action	Direct Feedback						Indirect Feedback						Combined Feedback			
	User 1		User 2		User 3		User 4		User 5		User 6		User 7		User 8	
	n	Q-value	n	Q-value	n	Q-value	n	Q-value	n	Q-value	n	Q-value	n	Q-value	n	Q-value
Music - English pop	14	4.79 (4)	12	3.79 (4)	15	4.06 (4)	15	4.41 (4)	8	3.61 (5)	3	**3.41 (4)**	22	2.83 (4)	9	3.78 (3)
Music - English rock	11	4.56 (4)	4	3.05 (1)	6	4.12 (3)	26	3.4 (5)	11	3.65 (3)	8	4.31 (5)	17	3.42 (3)	44	4.55 (4)
Music - Latin	5	3.99 (2)	21	3.81 (4)	7	4.96 (5)	8	2.08 (2)	10	3.43 (5)	1	**5.0 (2)**	10	3.98 (3)	11	1.16 (2)
Music - Noise	3	**2.69 (1)**	6	1.9 (1)	2	**3.25 (1)**	2	**3.25 (1)**	7	1.26 (1)	2	**3.69 (1)**	13	0.91 (1)	12	0.87 (1)
Music - Spanish pop	6	4.34 (3)	28	3.95 (4)	6	2.81 (3)	15	3.58 (4)	11	3.89 (5)	6	3.09 (4)	11	3.78 (4)	11	3.41 (3)
Music - Spanish rock	2	**4.56 (4)**	4	4.32 (1)	2	**3.94 (3)**	10	3.86 (4)	15	3.59 (4)	12	2.19 (5)	13	2.5 (3)	15	4.11 (4)
Photos - Animals	13	4.04 (3)	20	3.66 (4)	14	3.0 (4)	26	4.29 (5)	22	4.19 (5)	14	4.25 (4)	22	4.05 (4)	19	3.97 (4)
Photos - Landscapes	8	3.56 (4)	21	3.96 (4)	8	3.31 (4)	25	4.45 (5)	15	4.55 (5)	20	3.76 (5)	24	4.13 (5)	19	4.69 (5)
Photos - Monuments	7	3.48 (2)	13	3.75 (4)	5	3.47 (3)	9	3.15 (3)	8	3.85 (3)	10	3.61 (3)	12	3.95 (3)	30	3.49 (4)
Photos - Sad moments	8	1.4 (1)	10	**1.23 (1)**	11	1.17 (1)	8	1.59 (2)	6	1.71 (1)	9	1.27 (2)	12	1.27 (1)	21	1.33 (1)
Videos - Cooking	8	3.77 (4)	20	3.98 (4)	7	3.82 (3)	21	4.37 (5)	13	3.68 (5)	10	3.63 (4)	29	4.14 (4)	24	4.18 (4)
Videos - Sport	5	2.9 (1)	5	3.47 (2)	12	4.39 (4)	23	**1.15 (2)**	7	2.39 (1)	6	1.39 (1)	17	3.05 (3)	8	2.86 (2)
Videos - Film trailers	8	3.61 (4)	9	3.46 (3)	14	4.52 (4)	17	4.06 (4)	6	2.37 (3)	13	4.05 (4)	30	4.44 (5)	26	**4.1 (3)**
Videos - Funny videos	14	4.2 (5)	17	3.57 (3)	10	3.91 (5)	16	3.93 (5)	9	4.54 (5)	16	4.07 (4)	20	3.06 (3)	21	3.95 (3)
Videos - Comedy	4	4.25 (3)	5	3.63 (1)	7	3.53 (3)	7	3.35 (3)	8	3.23 (2)	16	**3.99 (4)**	14	4.07 (3)	12	3.84 (3)
Sessions completed (20 min)	15		20		15		20		16		16		16		16	

Table 9.4: Representation, for each participant and each entertaining activity, of the number of updates, final score, and initial score set by the participant of the adapted Q-values during the experiment. In blue, those Q-values whose number of updates are below 3. In red, the four representative case described in the previous section.

The results introduced in the previous table show differences when comparing the three conditions assessed in the experiment (direct feedback, indirect feedback estimated from the interaction metrics, and combining both types of feedback).

Figure 9.17 represents the activity exploration (the homogeneity in executing all the activities a similar number of times) in terms of the Relative Standard Deviation (RSD). As the results show, the robot explores more activities in the condition only using direct feedback than in conditions using indirect feedback and combined feedback. In the combined feedback condition, the number of updates of each activity is subtly higher than in the two other conditions. However, the gap between the most selected activities in this condition is more significant since combining both feedbacks fosters that the number of updates rises faster, provoking more considerable differences.

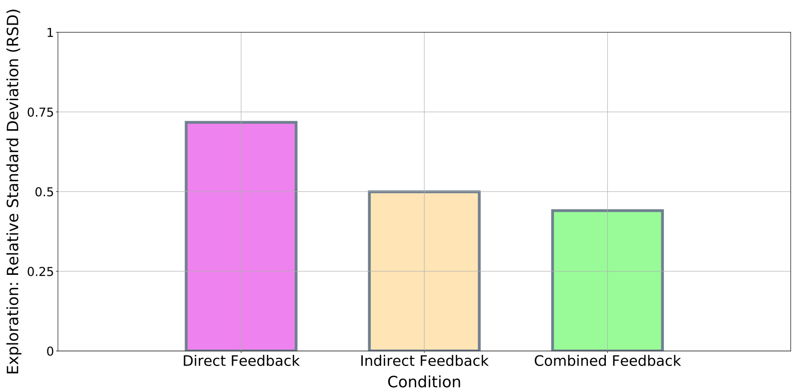


Figure 9.17: Normalised Relative Standard Deviation (RSD) for each condition tested in the experiment.

Finally, the Root Mean Squared Error (RMSE) computes the error between the preferences initially indicated by each participant and the final adapted results attained by the learning system. As Figure 9.18 shows, the error is minor when combining both direct and indirect feedback. This metric suggests that combining both types of feedback adapts the final Q-values better than individually using direct or indirect feedback. Consequently, Figure 9.18 supports our initial hypothesis about the best outcomes provided by combining both direct and indirect feedback.

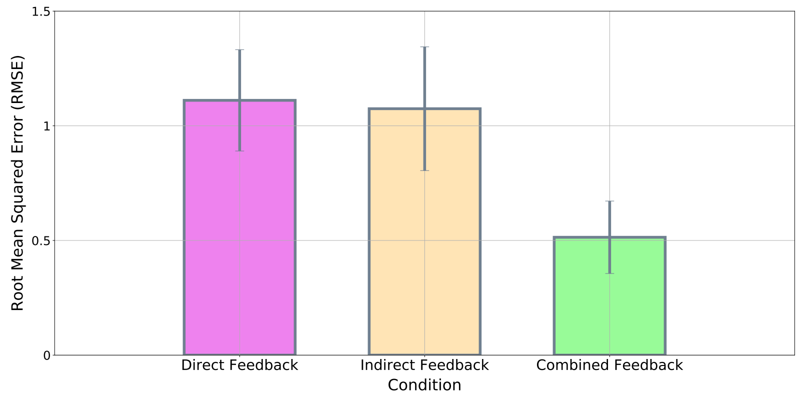


Figure 9.18: Root Mean Squared Error (RMSE) obtained in each condition of the experiment.

9.3.5. Discussion

The results of this third experiment provided very positive results in the system's operability. According to the experiment participants, the robot presenting adaptive mechanisms could suggest their favourite entertaining activities much more appropriately than a robot not presenting this adaptive behaviour. Besides, the interaction evaluation using the Godspeed questionnaire showed that the participants perceived the personalised robot as more likeable and intelligent than the general robot. These results prove the robot's correct operation in estimating the user's preferences, even though further experiments with more participants should be conducted to assess whether these preliminary results are also positive during long-term interactions.

This experiment was precisely oriented to anticipate the previous issue. During long-term interactions, the initial preferences of the user can variate. Besides, the predictions made by the system can not, in some cases, be as appropriate as expected. The system was improved by including an adaptive algorithm based on RL. This algorithm allows the robot to perceive and anticipate changes in each user's preferences. Using the interaction procedure that our robot has, we had two possible alternatives for obtaining the user impressions about each activity: directly asking how much (s)he liked the activity (direct feedback) or estimating this value from the interaction metrics (indirect feedback). Initially, we did not know which method was the best for obtaining the user's feedback. We designed a third experiment that aimed at whether the adaptive process works and which feedback provided the best outcomes. As expected, the combination of both direct and indirect feedback was the best alternative, as it reduces the error to the real preferences of the user and updates the score of each activity faster than just using direct or indirect methods independently.

9.4. Conclusion

In this section, we have presented three different experiments for endowing the DMS with the capacity to suggest the potential users of the robot those entertaining activities that they like the most. The first experiment comprised the evaluation of the user profiling method for actively identifying the users and obtaining their features to personalise HRI. The second experiment consisted of, using the LR algorithm that best suits our application, developing a system that can predict the initial preferences of a new user using information from other users that previously interacted with the robot whose features and preferences are stored in a database. Finally, a third experiment pretended to evaluate an improvement of the initial estimations. Using the feedback provided by the user while interacting with the robot, the idea is to adapt the initial predictions to the real preferences of each user in real HRI.

As mentioned before, the combination of LR algorithms with an adaptive mechanism based on RL provides very positive results in learning the user preferences. However, the system has some limitations that are worthy of mentioning. In the first place, the initial preferences' estimation depends on the datasets used for such prediction. The size and definition of the dataset are critical for the success of the estimation. Second, as mentioned above, the higher the number of labels to rank, the lower the ranking correlation obtained. Finally, the estimation of the indirect feedback is very sensitive to the interaction metrics used. These issues drive us to keep working on the PL System because we consider the area of robot personalisation an essential functionality of social robots.

Chapter 10

Evaluating the robot's Behaviour-based Learning System

The biological functions emulated in the *Motivational model* (see Chapter 6) provide the DMS with information about the robot state and situation. Then, using Reinforcement Learning (RL), we developed a learning method that allows our robot to map these situations and the robot's state into specific behaviours. To evaluate this functionality, we developed two scenarios:

- **Q-learning scenario:** The first scenario consisted of Mini interacting with a user. In this approach, the number of states and action was small. The results obtained in this scenario were published in [200].
- **Dyna-Q+ scenario**: The inclusion of new states and actions led us to investigate about methods that increase the learning speed of the system since the previous scenario using Q-learning reported very long learning periods in real scenarios. The results obtained in this second scenario were published in [355].

Following, we describe in detail both scenarios and present the results obtained in each of them.

10.1. Q-learning scenario

This section describes the Q-learning scenario we designed for endowing the social robot Mini with the capacity to map situations to actions using Q-learning, a well-known RL algorithm.

10.1.1. Aim overview

The robot's *Motivational model* provides the DMS with information about the robot current state. We define the robot's state as a combination of its internal and external state, including both its biological deficits that drive motivation and the influence of external stimuli. This state represents the situation that the robot is experiencing, leading it to behave in a specific way. Consequently, the goal of the behaviour-based RL system we proposed in this chapter is to provide the robot with a mechanism to learn how to map its current state to the best action it can execute. In this approach, we define the best possible action in a specific state as the one which provides the best benefits for the agent (e.g. reducing a need or allowing the robot to obtain a resource). At the beginning, the robot does not know the effects of its actions, so it has to explore them using trial and error. Once it has gained enough experience and learnt correctly, it will be capable of executing the best action sequence (behaviour policy) guaranteeing and optimal well-being.

10.1.2. Defining the Q-learning environment

This section explains the design of our learning scenario and describes the robot's state-action space. As described in Section 4.3.1, a RL problem requires the definition of the states in which the robot can be and the set of action it can execute to shape the learning scenario.

The combination of the robot's states with the actions it can execute shapes our learning scenario. This combination results in a number of situations where the robot has to learn how to behave. In this approach, we define the robot's state as a combination of its internal state (S_{inner}) and its internal state (S_{ext}). On the one hand, the internal state (S_{inner}) is the robot's dominant motivation, which is the motivational state with the highest level of intensity among all possible motivational states. As previously stated in Chapter 6, motivational states reflect the internal physiological and psychological needs. On the other hand, the robot's external state (S_{ext}) can be defined as how the robot perceives the states of other objects and agents with influence on its well-being. Considering the previous definitions, we represent the robot's state as $S = S_{inner} \times S_{ext}$.

Additionally to the robot's state space (S), a RL problem requires the definition of the action space A. In this scenario, our social robot Mini can execute 3 behaviour that will be explained in the following sections. The combination of the robot's state-action space define the number of learning situations (s, a) that the robot has to deeply explore to build its knowledge and behave well in the environment. Mathematically, these possibilities can be expressed as $n(S, A) = S_{inner} \times S_{ext} \times A$.

Following, we define the robot's internal state (S_{inner}), its external state (S_{ext}), and its repertoire of actions (A).

10.1.2.1 Internal state

The definition of the robot's internal state (S_{inner}) starts with a couple of biological variables that linearly evolve with time. Note that this first setting used a simple version of the *Motivational model* presented in Chapter 6 since it was the first version we developed, and the neuroendocrine responses were not integrated yet. Considering this information, the model considered the biological processes that represent the *Interaction* and *Rest* of the robot. Thus, the idea was to combine periods of relaxation with periods of interaction, according to the evolution of both biological processes, as depicted in Table 10.1.

On the one hand, the *Interaction* process increases its value with time at a rate of 0.2 units per second, from 0 units to 90 units. This process links to the *Social* motivational state. On the other hand, the *Rest* biological process reflects the need of the robot to take a *Rest*. Its value decreases in 0.16 units per second from 100 units to 0 units.

In this initial work, the deficit (d_i) of each biological process links to a motivation. The intensity of each motivation is calculated using Equation 10.1, being influenced by the deficit (d_i) of its associated biological process and the intensity of the external stimuli. Thus, like the last column in Table 10.3 shows, the intensity of the user stimuli modulates the deficit of the interaction process. This fact means that if the user is absent, the deficit associated with the interaction variable will be shallow.

$$m_i(t) = d_i(t) + d_i(t) \cdot si_i(t) \tag{10.1}$$

Motivation	Biological process	Initial value	Lower limit	Upper limit	Ideal value	Variation rate	Stimuli
Social	Interaction	0	0	90	0	+0.2	User
Relax	Rest	0	0	100	100	-0.16	None

Table 10.1: Motivations and biological processes used in the Q-learning experiment for endowing the social robot Mini with action-based learning.

The robot's motivational states reveal the intensity of its internal needs. Therefore, if a motivational state has a high level of intensity, it usually means that the biological process associated with it has a significant deficit. The computation of the dominant motivation follows a winner-take-all approach that selects, in each time step, the motivation with the highest intensity level. For becoming dominant, a motivational state has to be above 20 units, a threshold used to avoid shallow motivational states to rule the robot's internal state. Thus, the robot's internal state (S_{inner}) is the motivational state with the highest level of intensity among all active ones. If any motivation is a candidate to be dominant, the robot's dominant motivation will be defined as *None*. Consequently, the robot has three possible dominant motivations: *Rest*, *Social*, and *None*.

10.1.2.2 External state

In this scenario, we designed the user as the only external stimuli with influence on the robot. Therefore, the user is an external stimulus that influences the robot's internal and external state. As Figure 10.1 shows, Mini can perceive the user in the following states:

- **Absent:** The robot is not perceiving the user.

- **Far:** The robot perceives the user but not close enough to interact.
- **Sitting:** The robot perceives the user is sitting to interact with it.
- **Standing up:** The robot perceives the user is leaving the area of interaction.
- **Near:** The robot perceives the user in the area of interaction.

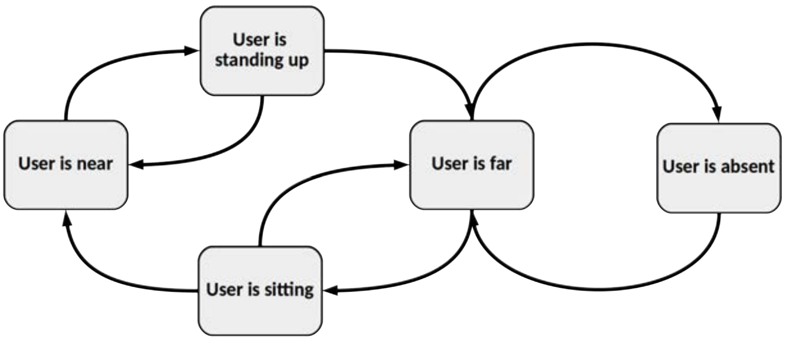


Figure 10.1: External state of the robot.

The previous state diagram not only shows the 5 possible user states, but the transitions among them. In this approach, the external state of the robot (S_{ext}) is exclusively defined by how it is perceiving the user. As mentioned above, it is important to emphasise that the external state of the robot is not the state of the user but the state of the robot concerning the state of the user, two definitions that mean different things.

10.1.2.3 Actions

Table 10.2 shows the set of behaviours of the social robot Mini and their effects on its biological processes during the experiment. Mini had 3 behaviours.

- The *Sleep* behaviour reduces the robot's fatigue
- the *Play game* behaviour reduces the interaction deficit
- The *Attract attention* behaviour is included to engage the user.

Behaviour	Effects
Sleep	Rest: −0.48
Attract attention	None
Play game	Interaction variable set to 0

Table 10.2: Set of behaviours and their effects of the social robot Mini during the Q-learning experiment.

The combination of robot's internal (S_{inner}) and external state (S_{ext}) define its global state (S). Using the previous definition, (S) is represented by the tuple (dominant motivation, perceived user state). If we add the set of robot behaviours, the tuple defining the state-action space (S, A) has the format (dominant motivation, perceived user state, action). In this experiment, since the robot has 3 possible dominant motivations, can perceive the user in 5 different states, and can execute 3 actions, number of situations the the robot has to learn is $n(S, A) = 3 \times 5 \times 3 = 45$. Since Q-learning is a tabular model-free method, each state-action pair (s, a) links to a unique Q-value ($Q(s, a)$) that is updated when the corresponding action finishes.

10.1.3. Experimental setup

Figure 10.2 shows the dynamics of the experiment we conducted to allow the social robot Mini learn how to behave and maintain an optimal well-being level. On the first place, the robot explored its environment in three sessions with a duration of 180, 90, and 60 minutes. During these sessions, the robot executed each action in each behaviour several times to acknowledge their effects. Then, in a exploitation session that lasted 40 minutes, we analysed if the robot correctly learnt the optimal policy of behaviour. In all sessions, the experimenter freely move in an empty room being perceived by the robot in one of the states represented in Figure 10.1. During the experiment, the robot autonomously selected the action according to the dynamics of the Q-learning algorithm.

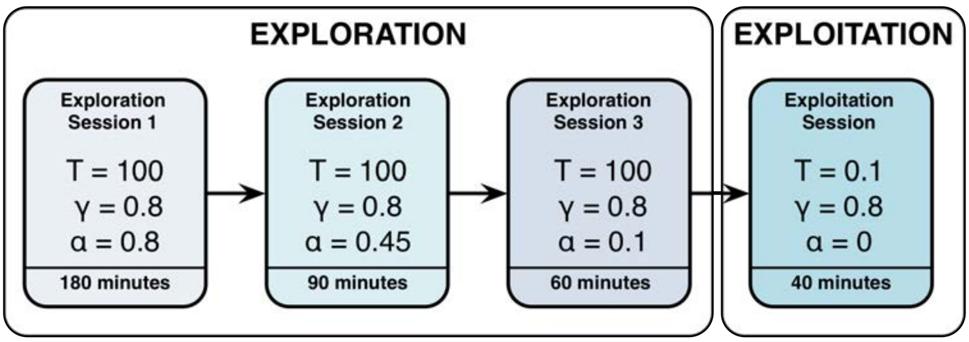


Figure 10.2: Stages of the experiment with Q-learning.

As Figure 10.2 shows, the set up of each sessions suffered significant variations that affected the updating of the Q-values using the Q-learning formula represented in Equation 10.2.

$$Q(s, a) \leftarrow Q(s, a) + \alpha(r + \gamma \max_a Q(s', A) - Q(s, a)) \quad (10.2)$$

In the first place, the discount factor γ of the Q-learning equation was empirically set to 0.8 units, balancing between future and current rewards. In the second place, the learning rate α decreased with time to give more importance to past experiences in contrast to new information.

Another critical parameter in updating the Q-values is the reward (r). Since the robot's goal was to optimise its well-being, drawing on previous studies [326, 327, 356, 357], we defined the reward value as the variation of robot well-being before and after executing an action. This value is represented by Equation 10.3,

$$r = \Delta Wb = \text{Wb(t=action finished)} - \text{Wb(t=action started)} \tag{10.3}$$

The well-being value (Wb), that ranges from 0 to 100, depends on the deficits of such biological processes can be calculated as:

$$\text{Wellbeing (Wb)}(t) = 100 - \frac{1}{M}\sum_{i=0}^{M} d_i(t) \tag{10.4}$$

where M is the number of biological functions and d_i their associated deficits.

During the experiment, the agent chose actions using the well-known Boltzmann method (see Appendix B.1 for further information). This probabilistic method, represented in Equation 10.5, assigns a higher selection probability ($p(a)$) to those actions with a higher Q-value in a particular robot state.

$$p(a) = \frac{e^{\frac{Q(a)}{\tau}}}{\sum_{b=1}^{N} e^{\frac{Q(b)}{\tau}}} \tag{10.5}$$

where $Q(a)$ is the Q-value of action a, τ is the temperature, and N is the number of actions the robot can execute.

The differences of the selection probabilities depend on the temperature T, indicating that T values close to 100 equal selection probabilities and low T values considerable differences in the selection probabilities. For this reason, during the exploration phase, T values were high to foster random action selection, and during the exploitation phase, T values were shallow 0.1, promoting actions with higher Q-values.

10.1.4. Evaluation

The evaluation of the learning scenario was carried out using two metrics:

- **Policy of behaviour learnt:** The optimal policy is the sequence of actions that the agent has to take in each state to fulfil the goal represented throughout the reward function.
- **Evolution of the well-being value:** The well-being value represents how optimal the internal state of the robot is (how reduced its deficits are). High well-being values (close to 100) mean an optimal internal state while low values (close to 0)

indicate a deficient internal state. Hypothetically, during the exploration phase, the well-being value should drop to not optimal levels. Nevertheless, if the robot correctly learns the behaviour policy, the well-being will present good values in the exploration phase.

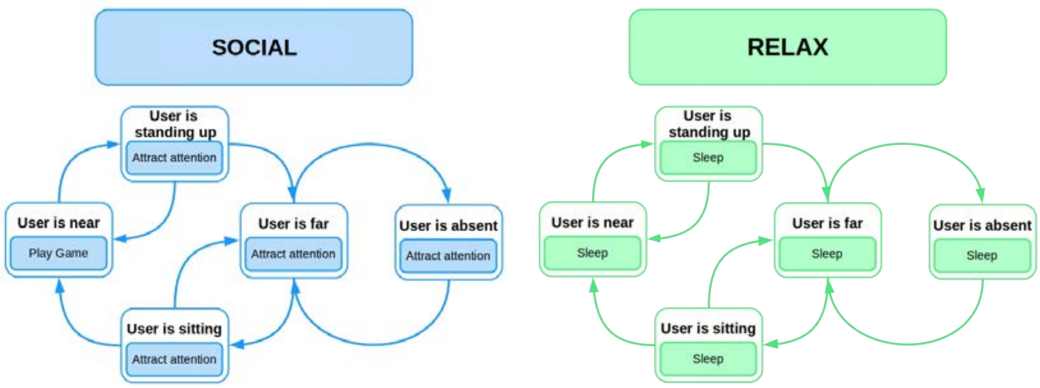
10.1.5. Results

The execution of the learning trials yielded interesting results that the following sections analyse.

10.1.5.1 Policy of behaviour learnt

Figure 10.3 shows the policy of behaviour learnt by the Q-learning algorithm in each robot state.

- When the dominant motivation is *Social* (see Figure 10.3a), the robot learns a sequence of behaviours to interact with him/her successfully. For example, when the user is absent, far, sitting or standing up, the robot learns that the best action to attract his/her attention is to get the user near and start the interaction. Once the user is near the robot, the interaction is possible, and the best alternative is to play a game. The interaction deficit will be reduced when playing with the user, ameliorating the robot's internal state.
- On the other hand, when the dominant motivation is *Relax* (see Figure 10.3b), the best possible action in all the possible external states of the robot is to sleep. *Sleeping* will reduce the rest deficit diminishing its level of fatigue and improving its internal state and well-being.



(a) Social dominant motivation. (b) Relax dominant motivation.

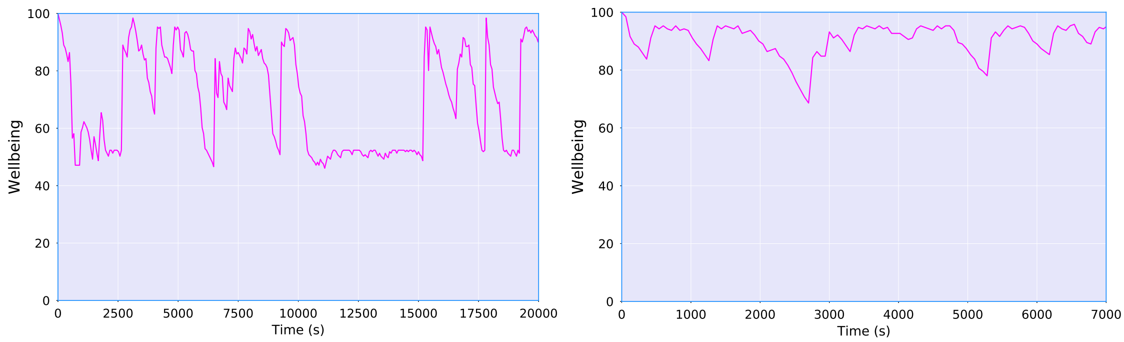
Figure 10.3: Policy learnt by Q-learning according to the robot's state (S). In each box we represent the user state perceived by the robot (S_{ext}) (top) and the learnt behaviour (bottom) regarding the dominant motivations (S_{inner}) of Social (left) and Relax (right).

These results suggest that Q-learning successfully learns the optimal behaviour policy to maintain well-being during the exploitation phase, as the following section where the well-being value evolution is analysed.

10.1.5.2 Well-being

In this experiment, the main goal of the robot is to maintain a good internal state. This event means that the robot has to execute actions that reduce those deficits with a higher level of intensity. Hypothetically, during the exploration phase, the well-being should be low because action selection is primarily random. However, once the learning process has been completed, the exploitation should lead the robot to maintain well-being in optimal values if the behaviour policy is correct.

Figure 10.4 shows the well-being evolution during the exploration (left) and exploitation (right) stages. During the exploration of the environment (left), the well-being oscillates between 50% and 90% most of the time, reaching optimal values on a few occasions. However, exploiting the behaviour policy the robot learns leads it to maintain excellent well-being, with values above the 70% during the experiment.



(a) Exploration phase.

(b) Exploitation phase.

Figure 10.4: Well-being of the robot during the exploration and exploitation phases.

10.2. Dyna-Q+ scenario

This section described the second scenario, in which we used a model-based Dyna-Q+ algorithm to improve the learning performance of the social robot Mini.

10.2.1. Aim overview

In a second approach, we improved our previous scenario [200] by using the Dyna-Q+ algorithm, a model-based RL method that we expect improves the performance of Q-learning. In this second stage, we increased the number of motivations, external stimuli, and actions of our initial proposal to improve the possibilities of our robot, defining a larger state-action space (S, A).

In this context, the results previously attained by Q-learning reported intractable learning periods. Besides, the stochastic rewards distribution and the long-lasting duration of Mini's actions put forward the necessity to use faster and stable alternatives. For this reason, we designed a second scenario to compare Dyna-Q+ with Q-learning and Q(λ) to evaluate the benefits of using a model-based method. In this setting, the goal of our learning method was the same as in the previous approach, allow Mini to learn how to behave for maintaining the best possible internal condition.

10.2.2. Defining the Dyna-Q+ environment

The second experiment for endowing the social robot Mini with action-based learning capabilities followed a similar dynamic to the previous one. The most significant difference resides in the use of Dyna-Q+, a model-based RL method to speed up the learning process.

Like occurred in the previous experiment using Q-learning, in this setting, we did not use the final setup of the *Motivational model* defined in Chapter 6. Instead, we define a more straightforward set of biological processes and motivational states that allow the robot to behave in a dynamic world. Although the scenario is not as complete as the shown in Chapter 2, it extends the one presented in the Q-learning approach.

The definition of the learning scenario follows the same one as described for the previous case. The robot's state (S) is a combination of its internal state (S_{inner}) and its external state (S_{ext}), mathematically represented as $S = S_{inner} \times S_{ext}$. On the one hand, the internal state (S_{inner}) is defined by the dominant motivation, and the external state (S_{ext}) by the states of external stimuli with influence on the robot. Following, we detail the definition of such states and the actions for this second scenario.

10.2.2.1 Internal state

Table 10.3 shows the biological processes of the social robot Mini in this second experiment. As the table shows, each biological variable links to a motivational state, like the previous experiment. The evolution of each biological process starts at an initial value, has lower and upper limits, and has an ideal value that represents the optimal value of the process. Besides, they follow a linear evolution that, in normal conditions, the robot's internal state deviates the value of each biological process from its ideal value producing a deficit (d_i) unless the robot executes specific actions. In this scenario, the robot can:

- Get tired, something represented by the biological process *Tiredness*.
- Get bored wanting to play, represented by the process *Boredom*.
- Need to socialise, represented by the *Social* process.

- Have eagerness of knowledge, represented by the *Knowledge* process.

Each biological process is related to a motivational state that represents the urge to satisfy the deficit of the corresponding biological process. As in the previous experiment, motivational intensities are calculated using Equation 10.1. Then, dominant motivation is computed using a winner-take-all approach and defining the robot's internal state (S_{inner}).

Motivation	Biological Variable	Initial value	Lower limit	Upper limit	Ideal value	Variation rate	Stimuli
Rest	Tiredness	0	0	100	0	0.1	None
Play	Boredom	0	0	100	0	0.3	User present Music player on
Socialise	Social	100	0	100	100	-0.2	User present
Learn	Knowledge	100	0	100	100	-0.3	None

Table 10.3: Definition of the biological processes and motivational state in the Dyna-Q+ learning scenario.

10.2.2.2 External state

In the second scenario we have included a new object, the *music player*. This object influences how Mini behaves, specially when learning sequences of action to attain a specific goal (e.g. turn on the music player for dancing). Mini can perceive this object as on or off. Besides, the state space of the user has been reduced to present and absent. Consequently, the robot's external state (S_{ext}) is now defined as a combination of how Mini perceives the states of the user and the music player, represented by a tuple *(user state, music player state)*.

10.2.2.3 Actions

In this scenario, the robot can:

- **Sleep:** reducing its fatigue and recuperating energy.
- **Wait:** for new upcoming events to occur.
- **Dance:** entertain itself dancing different songs.
- **Play with the user:** play a quiz game with the user.
- **Talk alone:** talk out loud expressing its feelings.
- **Talk with the user** talk with the user about general topics.
- **Search for information:** on the internet about the last news to learn about different things.
- **Turn on the music player:** turns on a virtual music player for dancing.

Table 10.4 shows the set of behaviours of the robot and their effects. Each of these behaviour produces a different effect on the biological processes of the robot or modify the state of the external stimuli, so the idea is to learn the most appropriate sequence of behaviours to maintain a good internal biological state.

Behaviour	Effects	Behaviour	Effects
Sleep	Tiredness: -0.5 Boredom: -0.1 Social: +0.1 Knowledge: +0.1	Talk alone	Tiredness: +0.2 Boredom: None Social: +0.5 Knowledge: None
Wait	Tiredness: -0.25 Boredom: None Social: None Knowledge: None	Talk with user	Tiredness: +0.2 Boredom: -0.2 Social: +0.5 Knowledge: None
Dance	Tiredness: +0.3 Boredom: -0-5 Social: None Knowledge: None	Search information	Tiredness: +0.2 Boredom: None Social: None Knowledge: +0.5
Play game	Tiredness: +0.2 Boredom: -0.5 Social: +0.2 Knowledge: None	Turn on the music player	Tiredness: None Boredom: None Social: None Knowledge: None

Table 10.4: List of behaviours of the social robot Mini and their effects during the Dyna-Q+ experiment.

In this second scenario, the robot state (S) is slightly different from the previous scenario, being defined by the tuple $S = S_{inner} \times S_{ext} =$ (dominant motivation, user state, music player state. As previously stated, the dominant motivation (S_{inner}) is the motivation with the highest intensity level, and the external state S_{ext} is formed by the user state represents if the robot perceives the user present or absent, and the music player state is if the robot perceives the music player turned on or off. Adding the action dimension, the robot's state-action space (S) combines 5 possible dominant motivations, perceives 2 user states, perceives 2 music player states, and 8 actions, giving a combination of $n(S, A) = 5 \times 2 \times 2 \times 8 = 160$ combinations, each of them represented by a unique Q-value $Q(s, a)$ as Dyna-Q+ is tabular method. In this environment, the state-action space (S, A) of the robot has significantly increased since the Q-learning scenario, so Dyna-Q+ the benefits that Dyna-Q+ may provide are essential to speed up the learning of the robot. Thus, when the robot executes a new action a in state s, there is an state transition from s to s', that leads the robot to a new situation. Consequently, the robot's goal is to learn the appropriate action in each state to maintain its internal state in good condition.

10.2.3. Experimental setup and evaluation

The experiment we designed for this second scenario consisted on two experimenters interacting with the robot at will. The robot was deployed in a particular house with the robot placed in the office room. In this case, we developed a long-lasting unique session with a duration of five natural consecutive days. At the beginning of the trial, the robot selected its actions with a high degree of randomness to test their effects in all situations. Once it gained experience, the robot exploited the learnt behaviour to maintain its internal state in the best possible condition.

The Dyna-Q+ scenario combines the power of model-free tabular methods with a model of the environment that allows simulating the system's performance in the background. The real performance of the robot is called acting while simulating in the background is called planning. Under normal conditions, the robot is in a particular state s where it has to decide which action it will take a (in this case, using the Boltzmann action selection method described in Appendix B.1). Once the action a has been executed in-state s, the robot receives a numeric reward representing how optimal it was to take action a in-state s. For this experiment, we updated our initial reward function (Equation 10.3) by a new formula (Equation 10.6 that considers the result of the action and the current dominant motivation to maximise the returned in each situation.

$$r = \begin{cases} \frac{1}{t}\left(0.5 \cdot \Delta d_{mot} + 0.5 \cdot \sum_{i=0}^{M-1} \Delta d_i\right) & \text{if } a \text{ succedded} \\ -1 & \text{if } a \text{ failed} \end{cases} \tag{10.6}$$

As stated above, the execution of action a makes the robot move from state s to a new state s', provoking a state transition. Despite the Dyna architecture can be combined with all tabular RL methods, we opted for using Q(λ) instead of Q-learning because it learns faster. The system combines the robot acting in the world and a model working in the background. Initially, the environment model does not yield any outcome but retrieves information from the real experiences of the robot and store them. Once it has gained enough information, it starts planning, simulating interactions about the appropriateness of executing specific actions in fictional states selected at random.

In our specific application, the model stores the rewards obtained by the robot after executing an action in state s and the next state where the robot is s' after executing that action. Using this information, the model creates a probability distribution about the rewards and following states after executing action a in-state s. Then, while the robot is interacting with the real environment, it selects a fictional state s at random, decides which action to execute using the Boltzmann's action selection method and, using the probabilistic distribution, estimates the reward r, the next state s', and updates $Q(s, a)$. As mentioned previously, for the model to start working, retrieving at least 10 reward values and state transitions in a particular state-action pair is necessary. Otherwise, the model can not produce estimations.

The planning process can be repeated as many times as desired per real acting, bearing in mind that choosing a high number of planning steps may produce that the model does not represent the real world, causing incorrect learning. In this setting, the number of planning per real acting has been empirically set to 10. Besides, the discount factor γ is 0.8 units, and the learning rate α exponentially decreases with the number of updates of each state-action pair $N(s,a)$. Using this setup, convergence is assumed by the law of large numbers [358].

Figure 10.5 shows the Dyna-Q+ model designed in this work. On the left side, it is possible to see the real acting of the agent, whereas the right side shows the planning stage. It is worthy of mentioning that one of the advantages of this architecture is to adapt to new situations. Thus, when the system saves more than 100 rewards and 100 state transitions for a state-action pair, the first data stored is removed to be aware of the changes in the environment.

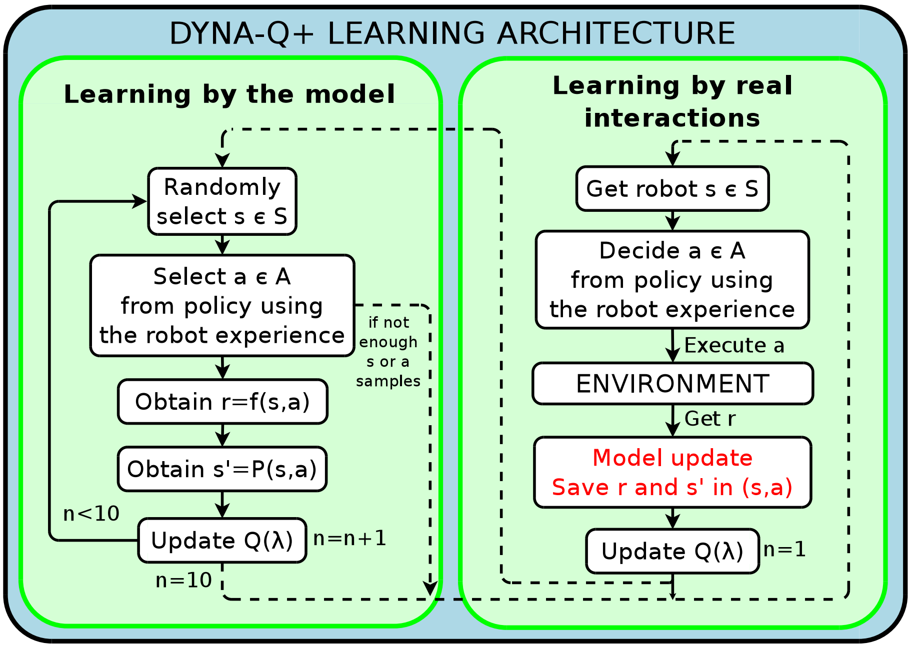


Figure 10.5: Dyna-Q+ architecture implemented in this thesis for speeding up the action learning of the social robot Mini. On the left side of the panel, the planning process is shown. On the right side, the real acting process.

Following a similar approximation to the first scenario, we evaluate our model considering that initially we had to compare if Dyna-Q+ outperformed classical model-free tabular methods. In this second scenario, we used the following metrics:

- For the comparison of Q-learning, Q(λ), and Dyna-Q+, we considered:
  - The real number of steps to convergence and the Mean Squared Error (MSE) of each algorithm. We assume that a Q-value has converged to its optimal value when the error between the final Q-value reached by the algorithm for a particular state-action pair does not exceed the 5% of error.

- The MSE to represent the error between the final Q-value learnt by the algorithms for each state-action pair ($Q(s,a)$) considering all the sequences of the intermediate Q-values learnt during the action exploration (see Appendix A.5 for further details).

- Once we found the best algorithm in our application, we used:
  - The optimal policy is the sequence of behaviours that the agent has to take in each state to fulfil the goal represented throughout the reward function.
  - The evolution of the well-being value before and after learning the best policy of behaviour.

10.2.4. Results

The following sections analyse the results provided by the learning system when the robot acted in the second scenario. First, we compare if the Dyna-Q+ model proposed in this work outperforms classic tabular methods. Then, we show the results in terms of the policy of behaviour learnt by the robot and the well-being evolution during the experiment.

10.2.4.1 Comparing the algorithms

In the second scenario, the algorithms Q-learning and Q(λ) were run simultaneously to obtain their real performance and compare the outcomes produced in each case. The initial goal of the experiment was to measure whether Dyna-Q+ overcomes model-free tabular methods in learning stability and speed. As Figure 10.6 shows, we measured these features in terms of the MSE and the average number of real steps to convergence. We emphasise the real number since Dyna-Q+ combines real experiences with simulated ones.

The results show that Dyna-Q+ yields better results in both cases than Q-learning and Q(λ). In terms of learning stability, the error of Dyna-Q+ is much smaller than the tabular methods. This outcome is due to the reward distribution obtained by each algorithm. While in model-free methods, the reward distribution is more sparse, using a model of the environment homogenises the reward distribution, making the learning process more stable. On the other hand, considering the real number of steps to convergence, plotted in Figure 10.6b, the comparison shows that while Dyna-Q+ requires around 52 steps on average, Q-learning and Q(λ) require more than 75.

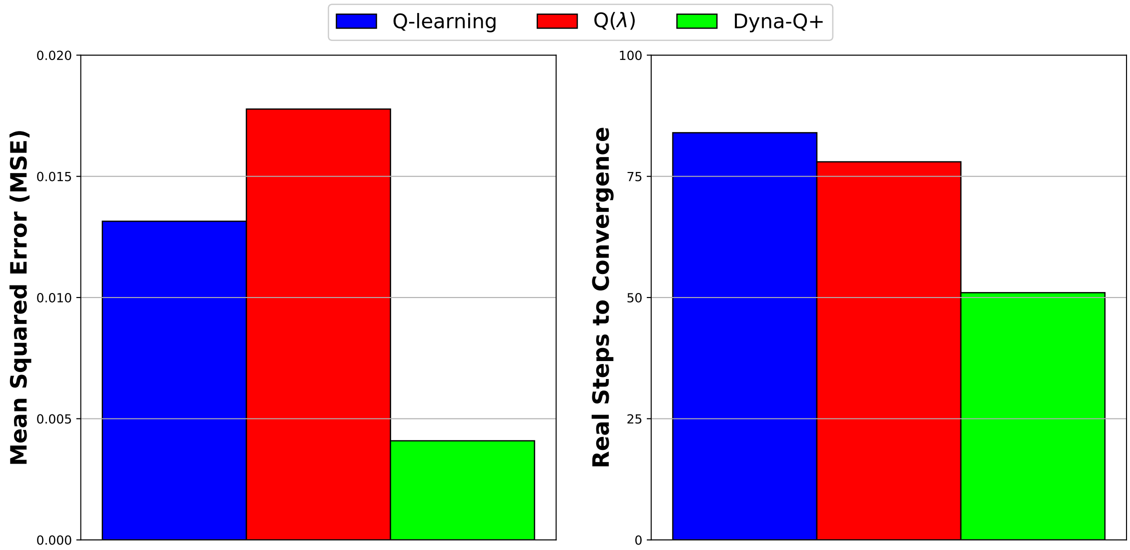
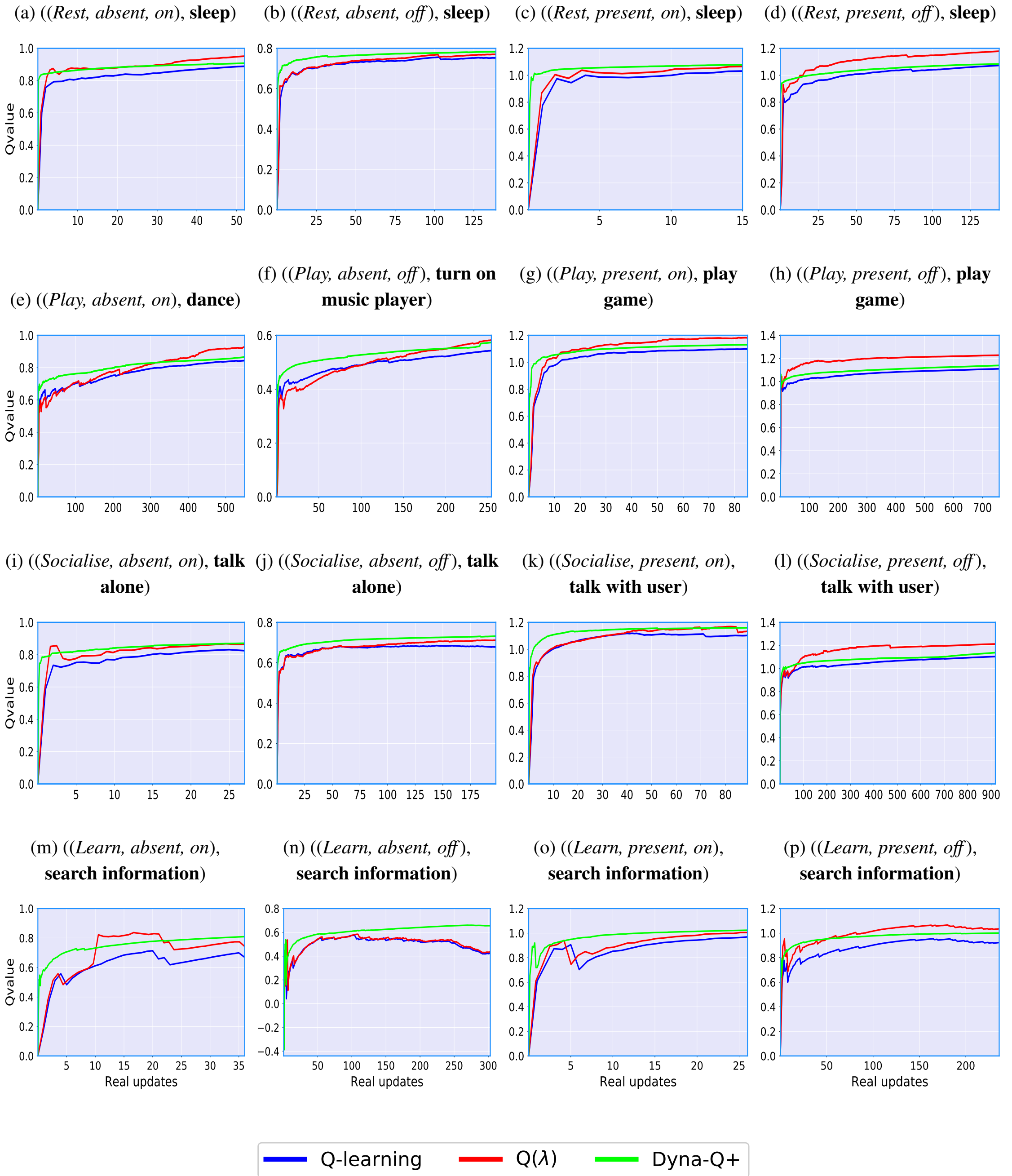


Figure 10.6: Average Mean Squared Error (MSE) (left-hand) and average number of real steps to convergence (right-hand) for each of the algorithms.

The previous comparison is reinforced by the graphs showing the most optimal Q-values for each robot state (see Figure 10.7). At a glance, it is possible to observe that the evolution of the Q-values produced by Dyna-Q+ converges faster and more stable than Q-learning and Q(λ). The graphs in Figure 10.7 show how each state-action pair has a different number of updates, proof that not all the robot's states are visited the same number of times. This issue ballasts the quality of the exploration, although the model used by Dyna-Q+ fosters the activation of the states that are not regularly visited in the real environment.



(a) ((*Rest, absent, on*), **sleep**) (b) ((*Rest, absent, off*), **sleep**) (c) ((*Rest, present, on*), **sleep**) (d) ((*Rest, present, off*), **sleep**)

(e) ((*Play, absent, on*), **dance**) (f) ((*Play, absent, off*), **turn on music player**) (g) ((*Play, present, on*), **play game**) (h) ((*Play, present, off*), **play game**)

(i) ((*Socialise, absent, on*), **talk alone**) (j) ((*Socialise, absent, off*), **talk alone**) (k) ((*Socialise, present, on*), **talk with user**) (l) ((*Socialise, present, off*), **talk with user**)

(m) ((*Learn, absent, on*), **search information**) (n) ((*Learn, absent, off*), **search information**) (o) ((*Learn, present, on*), **search information**) (p) ((*Learn, present, off*), **search information**)

Figure 10.7: Evolution of Q-values representing the best action for each robot state. In each graph, we compare the outcomes produced by Q-learning, Q(λ), and Dyna-Q+.* Note: The caption of each graph defines the state-action pair shown in the image, following the format (s, **a**) (e.g. ((*Learn, present, off*), **search information**)).

10.2.4.2 Policy of behaviour learnt

Once we assure that Dyna-Q+ outperformed the classical alternatives, we evaluated its individual results. Figure 10.8 shows the final Q-values learnt by Dyna-Q+ using a red to green gradient representation to highlight those which are optimal (green) above those that are not optimal (red). Besides, the Q-values that define the optimal policy of behaviour learnt by the robot has a frame in blue. The behaviour policy shows that when the robot the following outcomes for each motivation:

- When Mini is motivated to *Rest*, it will sleep independently of whether the user is present or absent or if the music player is on or off.
- If the robot is motivated to *Play*, it will turn on the music player if it is off and the user is absent. It dances if the music player is absent and the music player is on. It decides to play a game if the user is present.
- If the robot is motivated to *Socialise*, it will talk alone if the user is absent and talk with him/her if (s)he is present.
- Finally, if the robot is motivated to *Learn*, it will search for information on the internet to satisfy its knowledge.

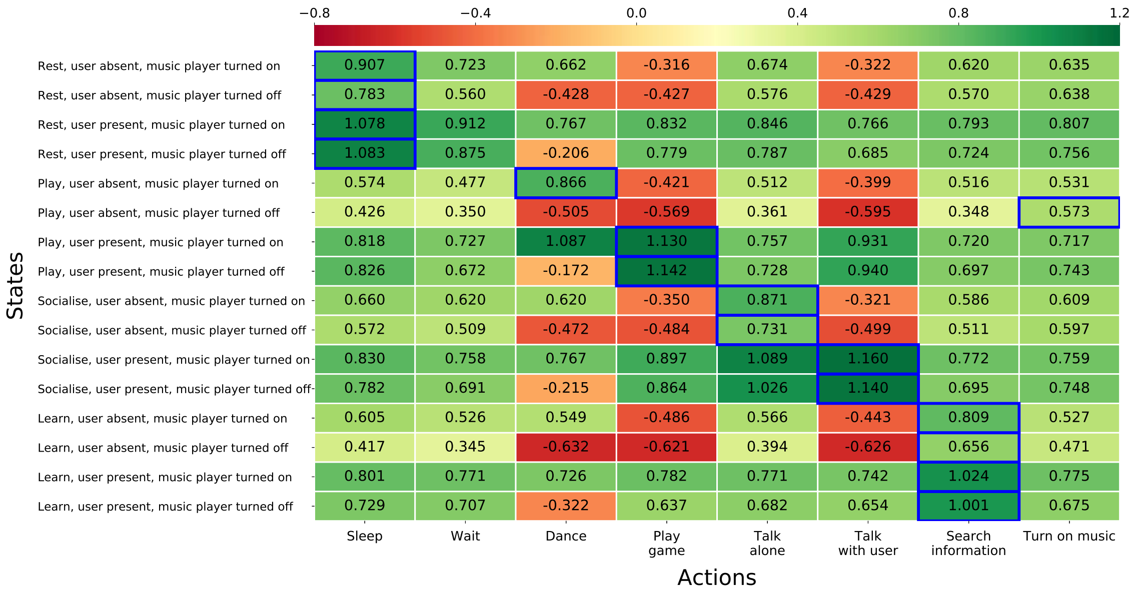
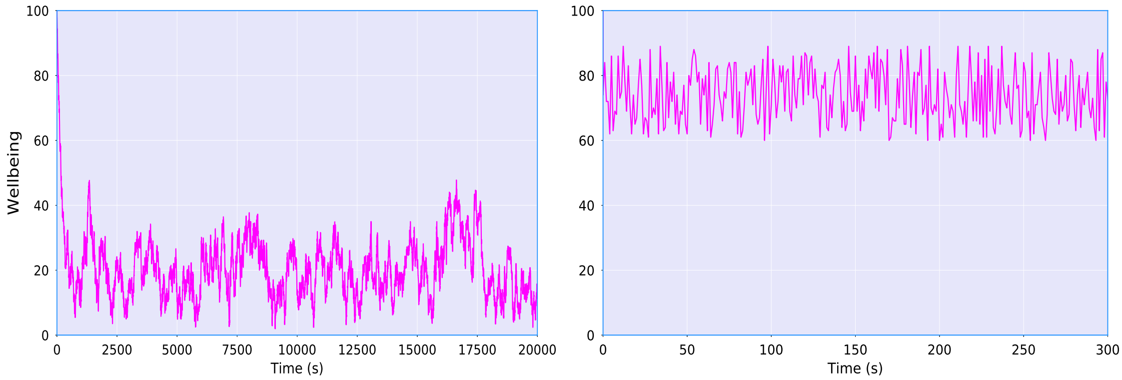


Figure 10.8: Q-values learnt by the Dyna-Q+ algorithm for each state-action pair. The optimal Q-values for each state-action pair are framed in blue. Actions with positive effects on the biological processes of the robot for each state are highlighted in green.

As the following section shows, the execution of the behaviour policy the robot learns leads it to maintain a good internal state.

10.2.4.3 Well-being

In this scenario the robot's well-being has a very similar pattern to the initial scenario only using Q-learning. During the exploration phase, the robot chose actions randomly, so the well-being of the robot dropped to levels below 40% in most of the trials. Since the robot correctly learnt the policy of behaviour that it has to execute, the well-being reaches values around 80% during all the experiments in the exploitation phase. Despite these positive learning outcomes, it is worthy to note that the well-being levels in the Q-learning scenario were much better than in the Dyna-Q+ experiment. In the second scenario, the number of biological processes increased with respect to the initial scenario using Q-learning, so satisfying four deficits is more difficult that just two. For this reason, in the second experiment is more difficult to maintain a high well-being level, showing that the definition of the environment has a strong influence on the performance of the learning system and the behaviour of the robot.



(a) Exploration phase.

(b) Exploitation phase.

Figure 10.9: Evolution of the robot's well-being in the exploration and exploitation phases of the Dyna-Q+ experiment.

10.3. Conclusion

In our effort of endowing our social robots with action-based learning capabilities, we designed two different environments. In the first one, a simple definition of the robot's state and using of Q-learning drove the robot to correctly learn how to behave to maintain an optimal level of well-being. However, the limitations of Q-learning, especially in terms of learning speed and stability when the state-action space is ample, took us to look for new algorithms to speed up the learning of the robot.

Finally, we opted to use Dyna-Q+, a model-based method that allows the robot to act and plan simultaneously. Among the benefits of Dyna-Q+, it is possible to stand out the speed up of the learning process for classical tabular methods like Q-learning, the homogenisation of the reward distribution, making more stable the Q-values fitting, the improvement of the environmental model perceiving unexpected changes and adapting to them, and the possibility to be combined with all model-free RL methods. However, Dyna-Q+ also presents multiple drawbacks that should be tackled in detail, like for example, the lack of generalisation since it requires a custom model of the environment, a notable increase of computational resources, or requiring non-tabular methods if the state-action space is considerable, something that increases the complexity of the system.

Like in occurs in all problems formulated by RL, the definition of the environment and the selection of the reward function is critical for the correct operation of the system. We have reached an optimal solution in these specific experiments, which supports the correct definition of the two environments presented above. Nevertheless, we believe that if the number of actions and external stimuli continues growing, function approximation and Deep Learning (DL) will be required to deal with large state-action dimensions. To conclude this discussion, we mention that action-based learning is essential for robots to deploy autonomous functionalities. However, from our point of view, the high number of hyperparameters necessary for tuning each case of study make their use in continuous applications very context-specific. For this reason, we consider it necessary to devise generalisation methods that extend the general capabilities of robots more than solve specific tasks.

CHAPTER 11

Evaluating the robot's biologically inspired behaviour

The following sections describe the validation of the *Motivational model* presented in Chapter 6.

- We begin in Section 11.1 with showing how the robot behaviour emerges due to circadian variations, biological functions, and external stimuli, using the *Motivational model* described in Chapter 6.
- Then, we present in Section 11.2 the results related to the social allostatic adaptive mechanism that allow Mini to develop a Pair-bonding (PB) with their users, variating its social behaviour depending on how each person socially behaves with the robot.

11.1. Motivated behaviour

In our architecture, the DMS selects, according to many different inputs, the most appropriate behaviour in each situation. One of the primary inputs is the dominant motivation, representing the robot's motivational state. We define the robot's motivated behaviour as the sequence of actions performed according to the robot's dominant motivation, aiming to maintain the robot's internal state in good condition. The previous section described two learning systems for endowing the social robot Mini with behaviour-based learning capabilities. Nevertheless, some related works in the literature do not focus on learning the optimal policy of behaviour but directly link specific behaviours to dominant motivational states. This approach is computationally easier for the robot since each motivational state directly executes a particular behaviour.

In this section, we present the last outcomes of the motivated behaviour exhibited by the social robot Mini, whose entire definition is presented in the *Motivational model* introduced in Chapter 6. Due to the large number of biological processes involved in generating the motivational states, we present the results divided into four phenomena.

- Sleep-wakefulness circadian signals.
- Affective response to environmental stimuli.
- Social interaction between the robot and the user.
- Stress response to threatening situations.

11.1.1. Mini's artificial biological processes

The *Motivational model* presented in Chapter 6 was carried out for its implementation in the social robot Mini. The idea is that the *Motivational model* sends periodic information to the DMS about the most important events and processes that are occurring in the artificial internal state of the robot. Then, the DMS makes autonomous decisions about which action to take in order to maintain the internal state of the robot in the best possible condition. In this approach, motivational states represent the current situation of fundamental biological processes occurring in the robot, like its needs, emotional responses, or reactions to threatening stimuli. As stated above, since the number of biological processes included in the Motivational model was huge to be tested altogether, we opted to assess them in four groups, depending on their influencing processes. Following, we describe the biological processes involved in each group and the behaviours that they elicit in the robot.

1. **Sleep-wakefulness cycle:** in our model, the light intensity, melatonin, and orexin affect the *Sleep* and *Wakefulness* biological functions. The deficits of such processes define our motivation to *Sleep* and *Stay awake*, urging the robot to *Sleep* and *Stay awake*.

2. **Affective response:** Mini can perceive many social stimuli that, acting with the hormones dopamine, serotonin, and brain norepinephrine define the robot's emotional reactions. Currently, Mini can experience *Anger*, *Surprise*, *Joy*, and *Sadness*, eliciting many reactive behaviours TO the *User presence*, and *Strokes* or *Hits* received.

3. **Social interaction:** The social stimuli perceived by Mini during the interaction (*User presence*, *Hits*, *Strokes*, *Fulfilling a goal* or *Not fulfilling a goal*) influence many hormones and neurotransmitters implicated in social behaviour (dopamine, brain norepinephrine, oxytocin, or arginine vasopressin). These substances define the *Entertainment* and *Social need* biological processes, whose deficits define the robot's motivation to Play with the user, *Socialise*, *Dance*, or *Request affect* (depending on if the user is present or not). Consequently, each motivation urges an specific behaviour (*Play quiz game*, *Talk*, *Dance*, *Request affect*) when it becomes dominant.

4. **Stress response:** Social stimuli and *Ambient noise* modify the robot's internal state rising the levels of stress hormones (CRH, arginine vasopressin, ACTH, cortisol, epinephrine, and norepinephrine). Abnormal levels of these substances increase the robot's *Stress*, a biological function that leads the robot to *Relax* in very stressful situation.

The stimuli and neuroendocrine substances influence many biological functions emulated in Mini. Then, these processes and their deficits define how the intensities of the motivations of the social robot Mini vary along the day. In a final step, the motivation with the highest intensity level becomes dominant and defines the robot behaviour. Thus, including such a biologically inspired artificial system provides the decision-maker with continuous information that allows Mini to deploy a long-lasting, natural, and expressive behaviour. The following section describes the experimental setup we designed to test the robot's biologically inspired behaviour during long-lasting periods.

11.1.2. Experimental setup and evaluation

The experiment conducted to assess the *Motivational model* performance consisted of the robot behaving according to its motivational states during four consecutive natural days (during day and night). During this period, the robot had to maintain its internal state in the best possible condition while reacting to unexpected stimuli perceived from the environment. An experimenter could interact with the robot by sitting in front of it. The experimenter could stroke and play a quiz game with the robot. Besides, although the experimenter principally exhibited friendly behaviour, the experimenter could mistreat it by hitting or refusing to play in some moments.

During the four days that lasted the experiment, we logged the values of all the processes running in the *Motivational model* for further processing and plotting their evolution. The evaluation of the system was carried out by displaying the most interesting events and situations occurring in the robot. As explained before, we focused on evaluating four interesting cases, each implicating different stimuli, biological functions, motivations, and behaviours. Then, we show the resulting behaviour exhibited by Mini during three days when combining all the biological functions while interacting with a user and when it is alone.

11.1.2.1 Sleep-wakefulness cycle

In this evaluation, we investigated whether the sleep-wakefulness cycle of the robot corresponded to the light-dark period during three consecutive days and adapted to subtle variations in the illumination level (light stimuli) perceived by the robot. Table 11.1 shows the processes involved in the sleep-wake cycle.

Stimuli	Neuroendocrine substances	Biological processes	Motivations	Behaviours
Light	Melatonin	Sleep	Sleep	Sleep
	Orexin	Wakefulness	Awake	Awake (not sleeping)

Table 11.1: Variables involved in the sleep-wake cycle.

11.1.2.2 Affective response

In this second study, we evaluated if Mini correctly affectively reacted to unexpected stimuli perceived from the environment during five hours where many different situations occurred. According to the definition of our *Motivational model* and the processed included in Table 11.2, the stimuli that Mini perceives modify the levels of dopamine, serotonin, and brain norepinephrine. Then, these substances affect the emotions emulated in the robot (Anger, Surprise, Joy, and Sadness), giving way to different affective behaviours.

Stimuli	Neuroendocrine substances	Biological processes	Motivations	Behaviours
User presence	Dopamine	Anger	React to user	User reaction
Stroke	Serotonin	Joy	React to stroke	Stroke reaction
Hit	Brain norepinephrine	Sadness	React to hits	Hit reaction
Goal fulfilled		Sadness		
Goal not fulfilled				

Table 11.2: Variables involved in Mini's affective response.

11.1.2.3 Social interaction

In this third case, we analysed how Mini's social behaviours changed depending on the stimuli it perceived from the environment and the levels of different hormones implicated in regulating social behaviour. As Table 11.3 shows, the *Entertained* and *Social nee*d biological functions derive in a wide range of social behaviours that vary depending on whether the user is present or not. This experiment was carried out during four consecutive hours where the social robot Mini and an experimenter interacted.

Stimuli	Neuroendocrine substances	Biological processes	Motivations	Behaviours
User presence	Dopamine	Entertained	Play with user	Play
Stroke	Oxytocin	Joy	Social need	Talk with people
Hit	Brain norepinephrine		Dance	Dancing
Goal fulfilled	Arginine vasopressin	Affect	Request affect	
Goal not fulfilled				

Table 11.3: Variables involved in Mini's social behaviour.

11.1.2.4 Stress response

Finally, we investigated how environmental stimuli can influence the stress levels in Mini. We tested Mini's stress response for almost four consecutive days in this scenario. As described in Chapter 6, we simulated an artificial human Hypothalamic-pituitary-adrenal (HPA) axis in Mini, intending to explore how the robot controls its stress executing relaxing behaviours. Table 11.4 shows the variables involved in regulating Mini's stress and the behaviours that the robot executes in stressful situations.

Stimuli	Neuroendocrine substances	Biological processes	Motivations	Behaviours
User presence	CRH	Stress	Relax	Meditate
Stroke	Arginine vasopressin			
Hit	ACTH			
Ambient noise	Cortisol			
	Epinephrine			
	Adrenal norepinephrine			

Table 11.4: Variables involved in Mini's stress response.

11.1.3. Results

The results obtained in the analysis of each group of biological functions studied in this dissertation are the following:

11.1.3.1 Sleep-wake cycle

The first results about the operation of the *Motivational model* comprise the sleep-wake cycle of the robot, represented in Figure 11.1. This figure contains the light stimuli in Figure 11.1 (a), the evolution of melatonin and orexin following circadian rhythms and affected by the light intensity in Figure 11.1 (b), the Sleep and Wakefulness biological processes in Figure 11.1 (c), the Sleep and Awake motivations in Figure 11.1 (d), and when the robot is sleeping and awake in Figure 11.1 (e). The inclusion of both processes pretends to endow the robot with relaxing periods, especially during the night, and periods of activity to be alert to new upcoming events (e.g. user's presence).

- As Figure 11.1 (a) shows, the sleep-wake circadian cycle starts with Mini perceiving the light intensity. This signal follows a circadian rhythm derived from the light-darkness period. In this case, we show the light evolution during three natural consecutive days.
- The light stimulates the secretion of orexin, a hormone implicated in maintaining the agent awake. Besides, the light inhibits the secretion of melatonin, a hormone that induces sleep. These facts are represented in Figure 11.1 (b), which shows the circadian evolution of these substances during the three days that lasted the

experiment. Since their circadian functions are cosine waves, in Figure 11.1 (b), it is possible to perceive such evolution. Melatonin and orexin's response to light varies depending on whether the light sensor captures natural light or artificial light. This effect appears in the light evolution. On the first and third days, the light signal represents the sensor's response to natural light, while the light spectrum of the second day corresponds to artificial light with a frequency of 50 Hz. Artificial light peaks during the midday of the second day represented in Figure 11.1 (a) provoking the inhibition of melatonin in Figure 11.1 (b).

- In the model, melatonin secretion increases the need to sleep in the robot. The higher the secretion rate of melatonin, the higher the speed with which the robot will need to sleep. Otherwise, orexin levels are indicative of the level of *Wakefulness* of the agent, like Figure 11.1 (c) shows. Acute artificial light makes the sleep deficit rise slower, provoking the robot to sleep subtly later on the second day than on the first day. It is fascinating how the light period influences the phase of the biological process and defines the wake-sleep cycle, synchronising the secretion of the hormones' circadian rhythms with the light phase to a limited extent. However, the circadian rhythms of both melatonin and orexin have a fixed pace independent of the light signal, establishing fixed hours where the robot should sleep.

- The deficit associated with the *Sleep* process increases the motivation to *Sleep* in the robot, while the levels of the *Wakefulness* biological process motivate the agent to stay *Awake*. Both biological processes act as opposing signals in charge of controlling the sleeping and awake periods of the robot.

- Figure 11.1 (e) shows that the robot sleeps when the motivation to *Sleep* is above 40 units, a threshold used to avoid the robot sleeping when the need is not very high. The execution of the *Sleep* behaviour reduces the levels of the biological sleep process, reducing the need for the robot to *Sleep*. When the robot is not sleeping, the robot maintains active since the *Awake* motivational state has a low activation threshold (5 units), serving as a base motivational state that activates when any other motivation is dominant.

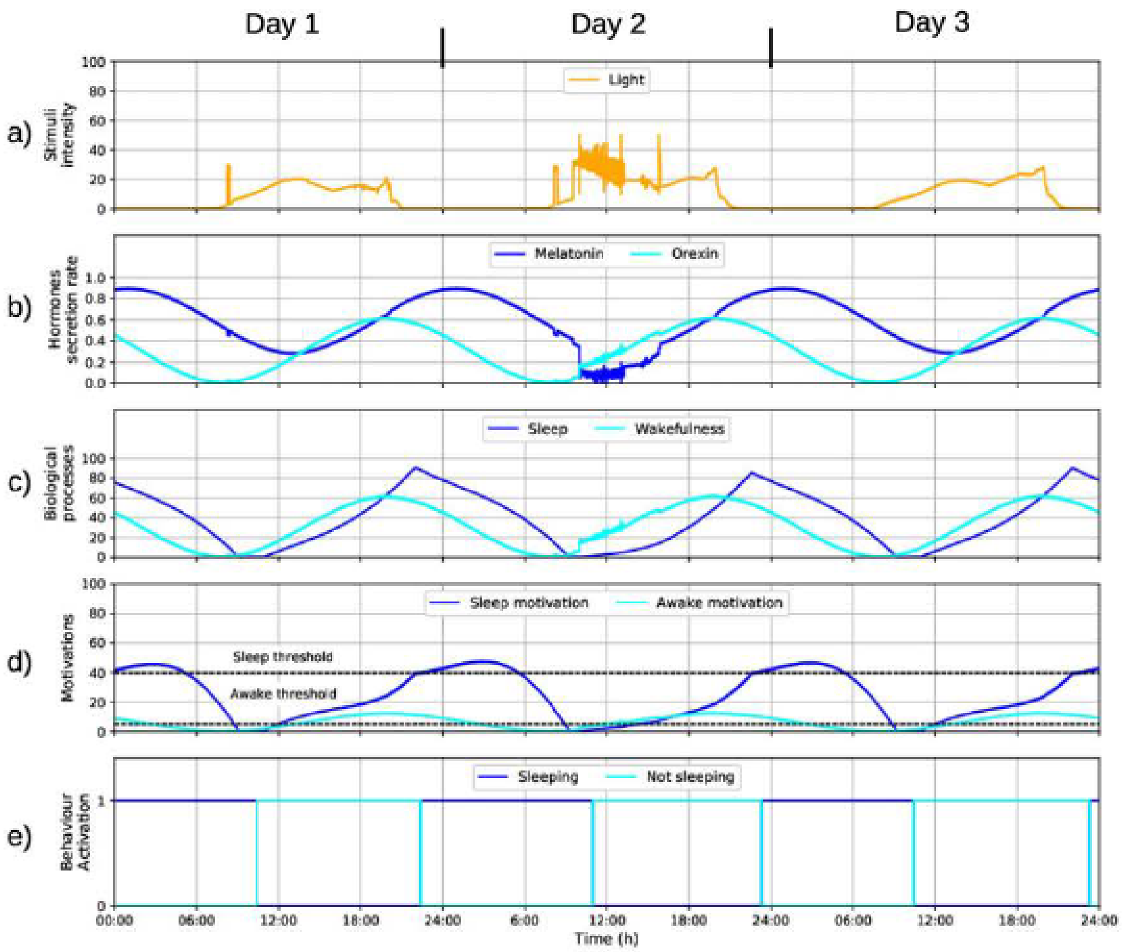


Figure 11.1: Circadian evolution of *Sleep* and *Wakefulness* biological process during three consecutive days. Both processes depend on the secretion levels of the hormones melatonin and orexin. The level of both hormones is highly dependent on light conditions. Then, their values define the intensity of the motivational states of *Sleep* and staying *Awake*. Once the motivational intensity is above its corresponding activation thresholds, it is a candidate for becoming dominant and defines the behaviour of the robot (in this case, *Sleep* or *Wait*).

11.1.3.2 Affective response

The second group of biological functions we have studied corresponds to the emotional reactions translated into reactive motivations. The processes involved in such responses are represented in Figure 11.2. Figure 11.2 (a) shows the stimuli with influence on Mini's affective state. Figure 11.2 (b) shows the dopamine, serotonin, and brain norepinephrine evolution, three substances that control Mini's emotions. Figure 11.2 (c) shows the values of *Anger*, *Surprise*, *Joy*, and *Sadness*, the four emotions included in our model. Finally, Figure 11.2 (d) depicts the emotional reactions that arise from the stimuli perception.

- In the model, the user presence, the strokes and hits, and the fulfilment or not the robot goals are stimuli influencing the chemical substances perceived by the robot, as shown in Figure 11.2 (a). The perception system of the robot provides the intensity of these stimuli.

- As Figure 11.2 (b) represents, the stimuli mentioned above affect dopamine, serotonin, and brain norepinephrine, the three most critical affective neuromodulators in the human brain. Besides, these substances are affected by their own circadian rhythm.

- As described in Section 6.5, Mini's emotions are *Anger*, *Joy*, *Sadness*, and *Surprise*. These emotions rise depending on the levels of the previous hormones, as Figure 11.2 (c) shows.

- Combining the emotional responses with the unexpected perception of certain stimuli, Mini can emotionally react to certain situations. Figure 11.2 (d) shows the response of the *Motivational model* in terms of reactive motivations when unexpected stimuli like the sudden appearance of the robot or the user hit the robot. Unlike in the sleep-wake cycle, in this case, Figure 11.2 concentrates on showing the middle hours of the third day because it was the period when the user interacted with the robot, leading to a wide variety of emotional responses. Positive stimuli like the user's presence, fulfilling a goal or being stroked evoke *Joy* or *Surprise* is specific circumstances.

  On the other hand, negative emotions like *Anger* or *Sadness* arise when the robot gets hit or when a goal is not fulfilled. It is worthy to note that for an emotion to be expressed in the robot, its level must be above 20 units to avoid mild emotional reactions that substantially affect the robot's expressiveness. Besides, like Figure 11.2 (c) shows, the intensity with which emotion is triggered depends on the perceived intensity of the stimulus that causes such reaction. Finally, it is vital to stand out that when reactive motivation triggers, it does with the maximum intensity level (100 units), making the behaviour of the robot change during a short time, showing the emotional reaction.

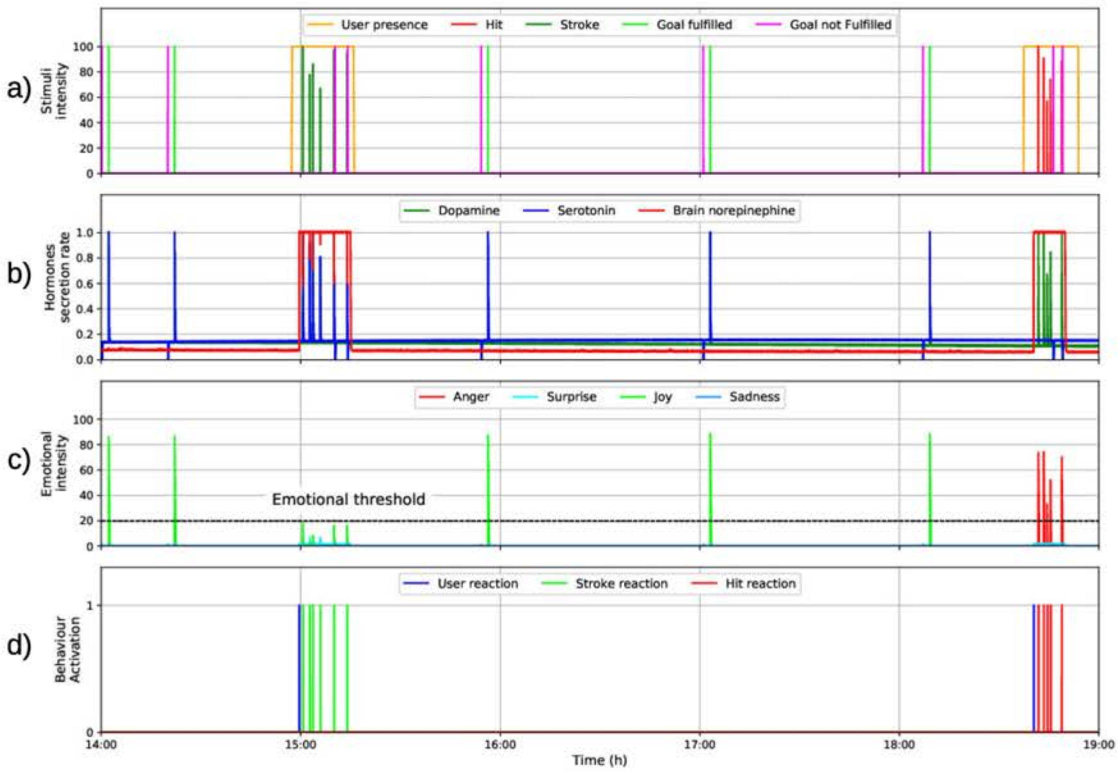


Figure 11.2: Emotional reactions as a response to dopamine, serotonin, and brain norepinephrine levels during five consecutive hours. The levels of the monoamine chemicals are strongly affected by the environmental stimuli perceived by the robot, causing emotional reactions in some circumstances. For an emotional reaction to being triggered, the intensity of the emotion must be above 20 units, a threshold empirically set to avoid eliciting soft emotions.

11.1.3.3 Social behaviour

One of the most important functions carried out by the *Motivational model* is the management of social stimuli. Although this functionality is extended in Section 11.2, Figure 11.3 represents small brushstrokes of the initial operation of the social interaction with the user. In this analysis, Figure 11.3 (a) shows the social stimuli that affect Mini's hormones related to social behaviour. Figure 11.3 shows the evolution of dopamine, brain norepinephrine, oxytocin, and arginine vasopressin, the four hormones that control Mini's social behaviour. Figure 11.3 (c) shows the biological processes that control Mini's *Entertainment* and *Social need*. Then, Figure 11.3 represents how the deficits in these biological processes shape the robot's motivational state, depending on whether the user is present or not. Finally, Figure 11.3 shows the behaviours that Mini's execute depending on the social situation it is experiencing.

- As in the previous section, Figure 11.3 shows the light hours of the third day, a period when the user interacted with the robot. In the model, the essential stimuli involved in social interaction are the presence of the user, the strokes and hits provided by them, and whether the tasks of the robot are completed or not (goal fulfilled and not fulfilled). The intensity with which the robot perceives these stimuli is represented in Figure 11.2 (a).

- These stimuli alter the levels of many substances in our model. As Figure 11.2 (b) shows, oxytocin and dopamine increase their levels with positive interaction like positive tactile contact and playing together, arginine vasopressin and brain norepinephrine are more involved in dealing with negative stimuli like receiving hits or being aroused by the user presence.

- Figure 11.3 (c) shows the evolution of the entertained and social need biological process. On the one hand, the *Entertained* process represents the urge of the robot for *Entertaining*, being affected by dopamine and brain norepinephrine. On the other hand, the *Social need* represents the necessity of affective contact being affected by oxytocin levels. The reduction of the *Entertained* deficit can be made by playing with the user if the user is present or by dancing if the user is absent. The *Social need* diminishes if the user strokes the robot or talks with it.

- The *Entertained* deficit translates into the motivation to *Play* and *Dance*, both modulated by the user's presence. The social need translates into the *Socialise* and *Request affect* motivational states, depending on whether the user is present or absent. Throughout the expression of both motivational states, the internal system controls the robot's social play and social needs.

- These motivational states map into different behaviours depending on which motivation is dominant. Thus, the robot will *Dance* when the dominant motivation is to *Dance*. When the dominant motivation is to *Play with the user*, the robot will launch the quiz game to play together. If the motivation is to *Socialise*, the robot and the user will talk. Finally, if the dominant motivation is to *Request affect*, the robot will attract the user's attention to request a stroke. All motivational states have their activation thresholds that avoid motivations with low intensities to become dominant. In these four cases, the thresholds are established in 30 units. Checking out Figure 11.3 (d and e), it is possible to appreciate that the *Request affect* and *Socialise* motivation never are dominant, mainly because the *Social need* deficit never reaches high values. This fact leads the robot never to Talk or Request affect from the user. The cause of this limitation is that the user provides affect without the robot requesting it, reducing the *Social need* and maintaining this deficit in optimal low values most of the time.

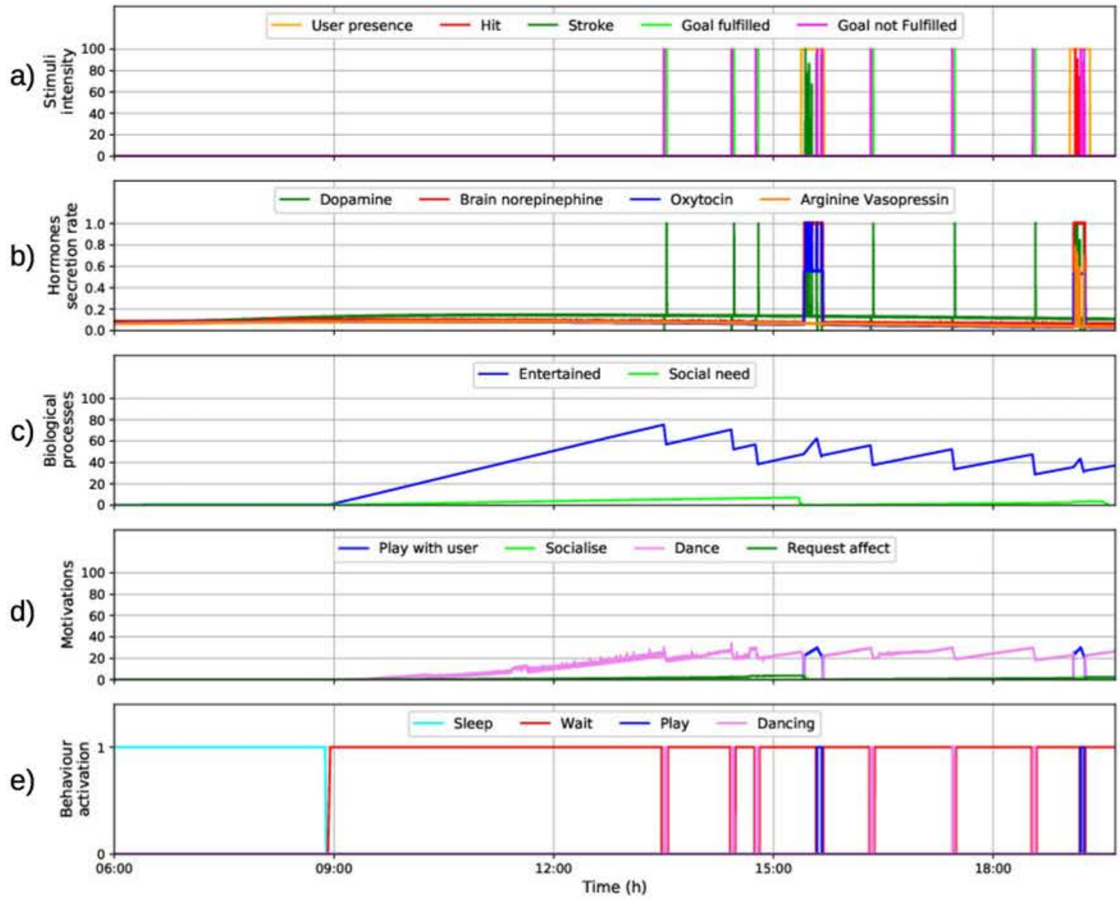


Figure 11.3: Social behaviour exhibited by the social robot Mini during interaction with a user. As the graphs show, social hormones react to the social cues perceived from the environment, stimulating the robot to interact with the user. The robot wants to *Play with the user* when positive cues are perceived. However, negative social interaction leads the robot never to execute the *Play* behaviour. In the absence of the user, the robot dances to reduce its *Entertainment* deficit.

11.1.3.4 Stress response

Finally, the last case represents how the robot manages the stress caused by threatening stimuli. Our model simulates a human Hypothalamic-pituitary-adrenal (HPA) axis in Mini for controlling its *Stress* response. Thus, Figure 11.4 (a) shows the stimuli that increase and decrease Mini's *Stress*. Figure 11.4 (b) shows the hormones implicated in regulating Mini's stress and Figure 11.4 (c) depicts the stress value during four consecutive days. This *Stress* level is translated into the motivation to *Relax*, that is represented in Figure 11.4 (d).

- As Figure 11.4 (a) shows, the stress hormones in our model are affected by several stimuli. Some increase hormonal levels related to stress, and others reduce them calming the robot. On the one hand, the ambient noise, the user presence, and the hits increase hormones related to stress, each in different quantities. On the other hand, when the robot receives a stroke, its stress is reduced, calming its internal state.

- In the model, we have characterised in Figure 11.4 (b) the evolution of the main hormones involved in the stress response: CRH, arginine vasopressin, ACTH, cortisol, adrenal epinephrine, and adrenal norepinephrine. These hormones stimulate the HPA axis response altering the behaviour of the robot when threatening stimuli are perceived.

- Figure 11.4 (c) shows the *Stress* level in the social robot Mini during three and a half consecutive days. The biological *Stress* process is directly linked to cortisol levels, which at the same time depends on the levels of other hormones like CRH, ACTH, and arginine vasopressin. These hormones increase their levels when aversive stimuli are perceived.

- According to the parameters set in our model, high ambient noise levels make stress levels increase significantly. Besides, the stress signal follows a precise soft circadian rhythm dependent on the circadian rhythm of the hormones in the HPA axis. Stress levels decrease if the user strokes the robot. Besides, the relax motivation increases with stress levels, eliciting the meditation if the *Relax* intensity is above 30 units. During the experiment, the stress levels never arose intensity levels for the *Relax* motivation to become dominant, excepting in the last hour of the experiment, when the user hit the robot, producing an essential increase of the stress and making the robot meditate for a while to reduce the stress levels. For this reason, we designed a particular experiment described in the following Section 11.2 to address this situation.

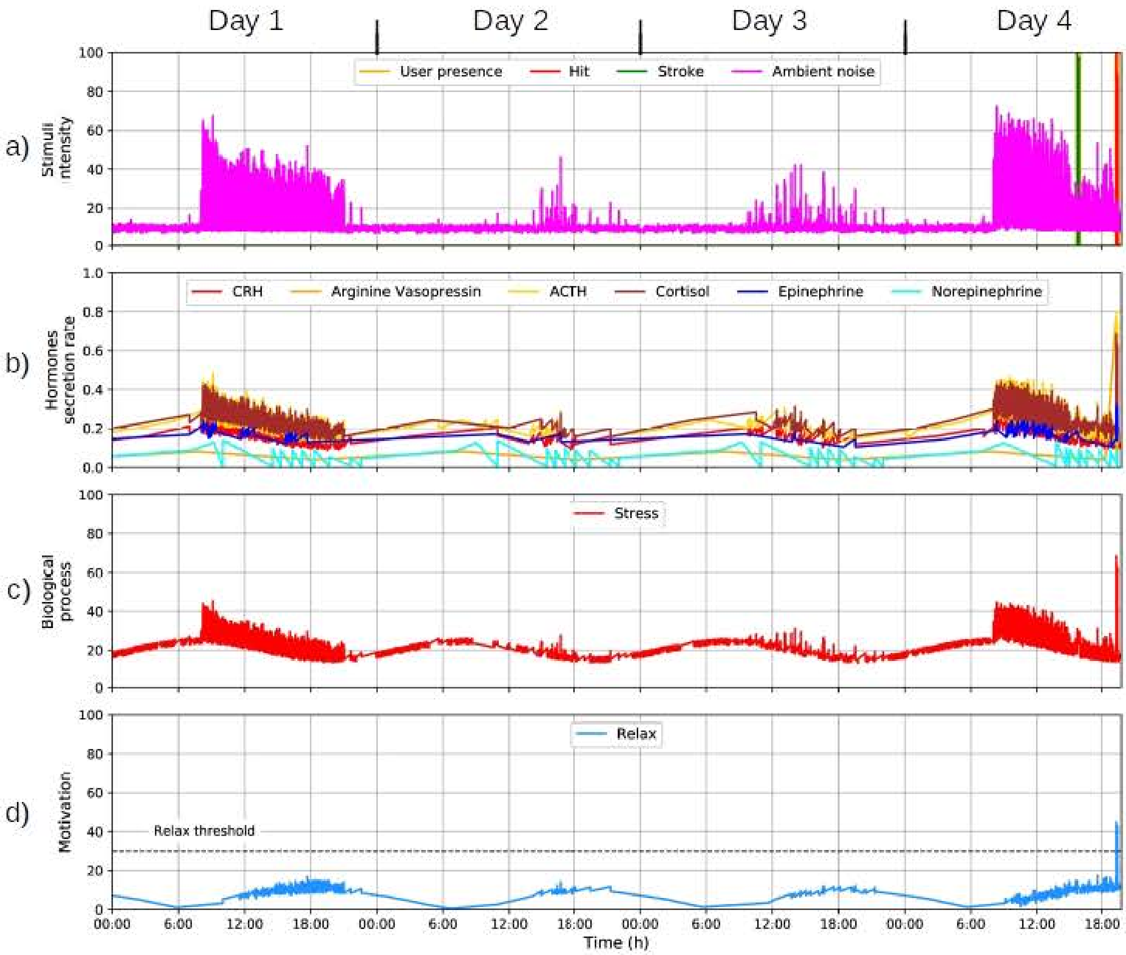


Figure 11.4: The stress response in the social robot Mini. The hormones in the HPA axis increase their levels with the perception of unexpected and threatening stimuli, mobilizing the body to avoid these aversive situations returning the internal organism to stability.

11.1.3.5 Resulting behaviour

The previous results individually analysed some biological functions emulated in Mini and the behaviour they trigger. In this section, we show Mini's long-lasting behaviour and the motivational variations when:

- Mini is alone and maintains its internal state in good condition.
- Mini is interacting with a user.

On the one hand, Figure 11.5 shows the variation of motivational intensities (a), the dominant motivation (b), and the behaviour executed by Mini (c) during three days. During the three days that are represented in Figure 11.5, it is possible to perceive that during the nights, the robot is motivated to *Sleep* so it executes the *Sleep* behaviour all night. During the daylight hours, the robot is awake and, once a while, dances to maintain its entertainment satiated. Since the robot is alone and does not interact with any user during this period, any motivational state that implicates the user (*Play*, *Socialise*, and *Request affect*) becomes dominant. Thus, when the user is absent, the robot mainly *Sleeps* during nights, *Dances* sometimes, and *Waits* for new events to occur while maintaining a good internal state.

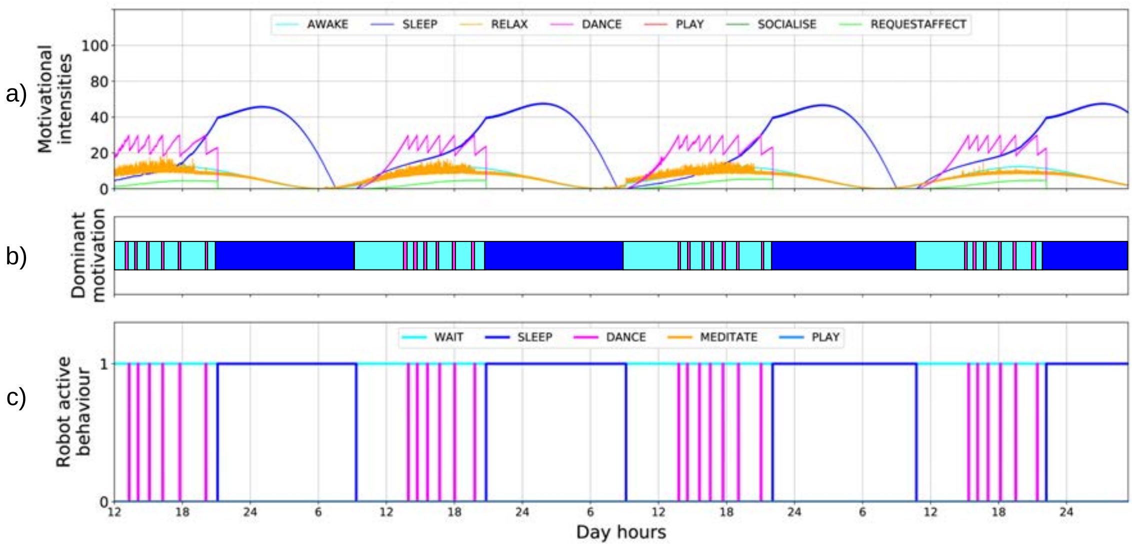


Figure 11.5: Long-lasting behaviour exhibited by Mini using the Motivational model when the user is absent.

On the other hand, Figure 11.6 shows the motivational intensities (a), the proactive motivations (top) and the dominant motivation (bottom) (b), and Mini's behaviour (c) while interacting with the user during five hours. As Figure 11.6 (b) shows, when the robot perceives certain stimuli (e.g. the user, strokes, and hits), proactive motivations like *Sleep* or *Play* combine with reactive ones like *Thank a stroke* or *Complain* by being hit. It is worth mentioning that one reaction can be executed with one long-lasting behaviour since reactions last a few seconds. Unlike the previous scenario, where the user was absent all time, in Figure 11.6 it is shown that when the user is absent but the robot needs to entertain, it *Dances*. However, when the user is present, the robot *Plays* with the user instead of *Dancing* alone fosters social behaviours. The rest of the time, the robot is *Waiting* for upcoming events.

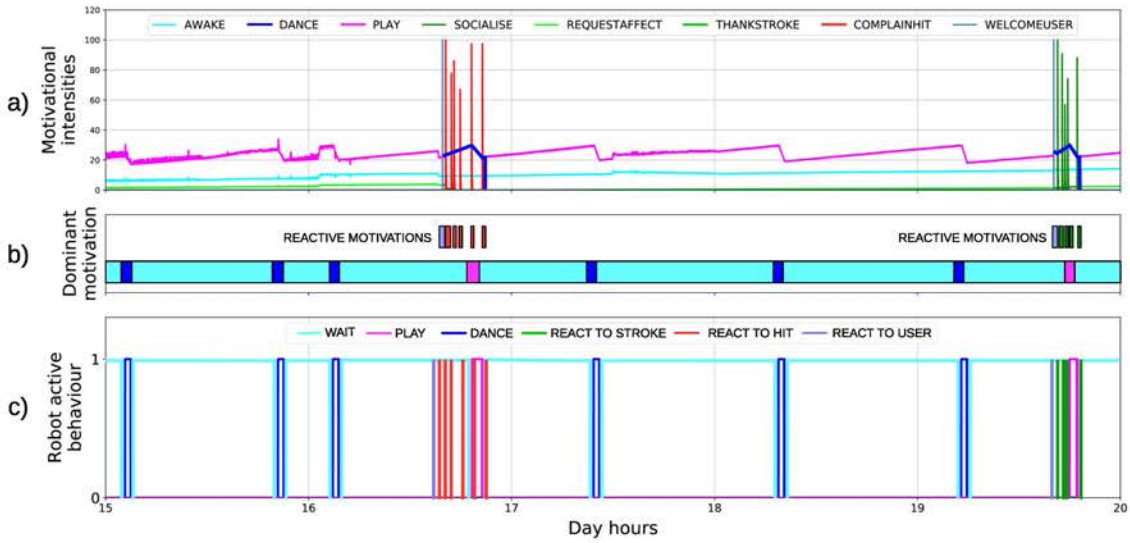


Figure 11.6: Behaviour exhibited by Mini using the Motivational model when interacting with the user.

11.1.4. Discussion

Although the high complexity in modelling biological processes for endowing the social robot Mini with a biologically inspired behaviour, we believe it is a promising area in researching artificial intelligence and social robots. Besides, the representation of neuroendocrine responses opens a broad range of opportunities for studying the effects of hormones and neurotransmitter secretion in the nature exhibited by artificial agents.

The primary contribution of the *Motivational model* introduced in this thesis is the inclusion of neuroendocrine responses to capture the effects of different stimuli in the organism, evolving following particular circadian patterns and promoting the appearance of deficits in the internal state of the robot, leading the agent to behave in a specific manner. Additionally, the model contemplates the combination and blending of proactive and reactive motivations, differentiating between long-term and reactive, short-term responses usually accompanied by affective charge. Unlike in the previous two chapters, the biologically inspired behaviour generated using neuroendocrine responses does not use learning. Instead, we propose adaptive mechanisms to stimuli and circadian rhythms. We evaluated four interesting cases that represent the robot's sleep-wake cycle, affective response, social interaction, and stress response during our experiment. Although we analysed them separately, all of them simultaneously run in Mini.

- Figure 11.1 showed how the robot has a sleep-wake cycle that coherently allows it to sleep in the dark period and stay awake during the light hours. Besides, due to the inclusion of circadian rhythms, light perception, and hormones, the sleeping-awake phase could autonomously adapt to the light hours without including any learning mechanisms or predefined information.

- The affect generation is shown in Figure 11.2 is another necessary functionality of the model, allowing the robot to express affective signals. In order to endow our robots with complex and reasonable affect generation, we pursuit to define a system that shows the time effect of affective responses in humans. For this reason, and like it is further detailed in Chapter 12, the affective component of the robot considers emotional and mood responses.

  On the one hand, emotions are triggered after perceiving arousing stimuli that stimulate monoamines (dopamine, serotonin, and brain norepinephrine), and finally, the appearance of emotions. The intensity of the stimuli is determinant to trigger emotions. Besides, the model deals with emotional blending, allowing the agent to experience different emotions simultaneously. When the arousing stimuli are no longer perceived, the intensity of the activated emotions decays with time. The accumulation of experiences defines the robot's mood, a stable affective state. From our point of view, the expression of affective cues is essential for natural communication between the user and the robot since it allows the user to know how the robot feels.

  Nevertheless, the affective component implemented in the *Motivational model* is just a simple representation of the affective generation in humans. For this reason, we are conscious that further research is necessary to continue enhancing the affective generation and expressiveness of social robots. Besides, user evaluation is fundamental as expressing affective cues in machines is not an easy task.

- Figure 11.3 showed the social behaviours exhibited by the robot in different situations. On the one hand, the robot behaves differently depending on if the user is present or not. Thus, if the user is present, the robot will typically foster an interaction rather than entertain executing the dancing behaviour. Besides, since the user spontaneously stroked the robot once a while, the *Social need* never presented very high levels to activate behaviours where the robot demands affect.

- Finally, we have shown in Figure 11.4 how the simulation of a HPA axis can manage the stress levels in the robot. However, our initial experiment did not show any high-stress situation that led the robot to execute relaxing behaviours and refuse to interact with the user. For this reason, the following section explores more deeply how stressful stimuli may lead the robot to refuse to interact.

To conclude this discussion, we would like to emphasise the future of modelling biological processes in social robots to attain a natural behaviour. In our opinion, just like autonomous systems requires a DMS, social robots require biological foundations to exhibit utterly natural behaviour. Although we developed the last version of the *Motivational model* in the final stages of the thesis, we are amazed by its great potential. As the future work section later in this manuscript describes, we have many ideas to continue expanding the functionalities of the *Motivational model* in order to endow the robot with new capabilities.

11.2. Human-robot pair-bonding

Social relationships are essential in humans. How we relate with other people impacts our internal state in many different ways. For this reason, the model presented in Chapter 6 emulates a biological process that we have called Pair-bonding (PB) that regulates the relationship between the robot and each user depending on the interaction process. To test the influence of the PB on the robot's social behaviour, we designed an experiment that shows the robot's response when facing different types of users.

11.2.1. Pair-bonding modelling

In our *Motivational model*, presented in Chapter 6, the effect of the external stimuli on the robot's internal state was modelled using Equation 11.1,

$$se_k(t) = \alpha_k \cdot \beta_k \cdot si_i(t) \tag{11.1}$$

In the previous equation, se_k is the influence of a stimulus on a robot's biological function, si_i is the intensity with which Mini perceives the stimuli, α_k is a fixed weight that modulates the physical impact of the stimulus, and β_k is a variable weight that modulates the psychological interpretation we made from stimuli.

According to our model, stimuli affect hormones, neurotransmitters and biological functions (e.g. light modulates pupil size). Consequently, we have designed a system where the weights α_k and β_k strongly influence how stimuli affect the robot's internal processes. More specifically, in this experiment, we designed an allostatic mechanism that regulates the cognitive interpretation (β) of social stimuli according to the PB with the user. Thus, the robot will adapt its behaviour depending on whether the relationship with the user it interacts with is good or bad. For example, if the robot maintains a positive relationship with a user but (s)he suddenly hits the robot, the robot will calm instead of refusing the interaction with the user. Otherwise, if the PB between both users is terrible, the robot will refuse to interact with the user without paying attention to them.

The deployment of such adaptive mechanisms starts by including a biological process in the robot named PB. This process is user exclusive, so after recognising the user, the robot will load the previous PB they had. The PB ranges from 0 to 100 units, meaning 100 an excellent relationship between both agents and 0 a null one. From scratch, the new users and the robot have a PB of 50 units, defining a normal relationship between them. The PB increases when the robot receives positive social stimuli from the user, like strokes, correct answers, or succeeding in the task both agents are performing. Otherwise, the PB deteriorates, especially if the user hits the robot or is not engaged with it. Time also significantly impacts the PB between the user and the robot since the PB value decreases if the communication does not exist for long periods. The modelling of the impact of each stimulus on the PB and the neuroendocrine substances involved in social behaviour is represented in Table 11.5.

Stimulus	Neuroendocrine substance	α_k	β_k
Friendly user presence	Oxytocin	+0.05	+0.02∗ PB
Stroke	Oxytocin	+0.05	+0.02∗ PB

Table 11.5: Effect of PB in how the robot perceives and interprets stimuli.

The cognitive interpretation (β) that the robot makes about social stimuli depends on the Pair-bonding (PB) between the robot and the agent. Since the PB ranges from [0, 100], but the value of β is limited to $[0.01, 2]$, the value of the PB maps to the range of β. Thus, if for example the PB is excellent and values 100, β's value will be 2. This adaptive perception of the stimuli makes that if the user-robot PB between both agents is good, the perception stimuli of social stimuli will be amplified, leading to a notable improvement in the social reaction and internal state of the robot, fostering social communication like playing together. However, if the PB presents values below 50 units, indicative of a bad relationship between both agents, the impact of social stimuli will be reduced, leading the robot to be more suspicious and, in extreme cases, refusing the interaction with the user. In this approach, we would like to remind that the PB value is calculated using Equation 11.2,

$$PB = PV + 0.2 * OT - 0.2 * AVP - 0.01 \tag{11.2}$$

where PV is the PB value in the previous time step, OT is the oxytocin level, AVP is the arginine vasopressin level, and the factor -0.01 reflects the deterioration of the human-robot relationship with time.

The following section deepens in the experiment carried out for testing the adaptive social system of the robot, focusing on enumerating the behaviour responses of the robot according to the evolution that the PB with the user takes.

11.2.2. Experimental setup and evaluation

Three different user profiles we considered for interacting with the social robot Mini.

1. A friendly user that positively interacts with the robot. This user always provides positive tactile contact by stroking the robot, responding to its questions, and playing with it.

2. An aversive user negatively interacts with the robot. This user never strokes the robot but continuously hit it. Besides, when the robot wants to play, the aversive user ignores the questions asked by the robot.

3. A neutral user that balances both previous approaches. The behaviour of the neutral user sporadically changes between the friendly and aversive user profiles.

The experiment consisted of a participant following each of the three profiles defined above. During these interactions, the experimenter was seated in an empty room in front of the robot. The duration of each user-robot interaction lasted around 25 minutes.

This experiment was focused on the evolution of the social relationships between the robot and the user. For this reason, we limited the biological processes and robot behaviours used in this scenario to the ones shown in Table 11.6. Considering this limitation, the robot could perceive the presence of the user and load the level of PB between both agents before starting the interaction. Besides, Mini could perceive strokes and hits from the user. Finally, Mini could differentiate between correct and wrong answers and the user disengagement. The robot had three artificial neuroendocrine responses directly involved in social behaviour: dopamine, oxytocin, and arginine vasopressin. The role of these chemicals was to respectively control the need to *Play*, *Socialise*, and *Escape* from aversive situations. As described in Chapter 2, dopamine is a hormone involved in social play behaviour, oxytocin controls prosocial behaviour increasing with positive social interactions, and arginine vasopressin regulates aggressive and avoiding reactions to threatening stimuli.

Stimuli	Neuroendocrine substances	Biological processes	Motivations	Behaviours
User presence	Oxytocin	PB	Relax	Meditate
Stroke	Arginine vasopressin	User rejection	Play	Talk
Hit	Dopamine	Wakefulness	Entertained	Socialise
Correct answers		Social need		Wait
Wrong answers				

Table 11.6: Variables involved in regulating Mini's PB and its social behaviour.

As Table 11.5 shows, additionally to the PB biological process, the internal state was characterised by the level of *Wakefulness* to maintain the robot active during the experiment, the level of *Entertainment* to represent the desire of playing with the user, the *Social need* of the robot, and the *User rejection* as an opposite signal to the PB. The evolution of these processes depended on the secretion rates of the hormones of the robot, as shown in Table 6.5. Then, the intensity and deficits of these processes define the activation of the motivational states that drive behaviour, as Table 6.7 shows. Thus, when the robot is motivated to *Play with the user*, it executes a quiz game. When the robot is motivated to *Socialise*, it talks to the user. In case it is motivated to *Relax*, it meditates. If the robot does not have any previous motivational urges, the *Awake* motivation is dominant, maintaining the robot awake but not executing any specific action. To improve both the naturalness and reactivity of the robot, we included two reactive motivations as a response to hits and strokes.

The representation of the *Motivational model* used in this experiment allows us to explore in-depth those processes directly involved in HRI from a biological perspective. The idea of endowing the system with the capacity to adapt its behaviour depending on how each user behaves with the robot provides a more natural and flexible behaviour, as the following results show.

The results obtained from the previous experiment were saved for representing the most important processes involved in regulating Mini's social behaviour. The log system updated the value of each process with a frequency of 0.5 seconds. Besides, the actions that the robot executed were saved to know which action the robot takes in every moment.

Once the experiments of the three conditions tested finished, we represented the results in three independent graphs showing the evolution of the essential processes in each robot-experimenter profile interaction. Each graph contains five-axis showing:

a) The PB evolution.

b) Hormonal values regulate the biological processes of the robot.

c) Intensity of each stimulus perceived by the robot.

d) Motivational states of the robot.

e) Mini's behaviours.

11.2.3. Results

The following sections show the results of the three scenarios presented above. The first section showed the evolution of the PB and its impact on its internal state and social perception when the robot faces a friendly user. In the second section, we show how the PB evolves when an aversive user interacts with the robot. Finally, the results of the social communication of a neutral robot interacting with the robot are shown.

11.2.3.1 Friendly user

The interaction with a friendly user yielded valuable results in terms of PB and the range of behaviours exhibited. As Figure 11.7 (a) shows, the PB between the robot and the user is moderate at the beginning of the experiment (around 50 units). Initially, the robot is not perceiving the user, so the robot decides to stay idle and wait for new incoming events. During the waiting period, the three hormones' secretion values in social behaviour maintained basal levels below 0.2 units per time step (see Figure 11.7 (b)). Around 90 seconds after the start of the experiment, the robot perceives the user is sitting in front of it. The mere presence of the user provokes an increase in the three hormones since it stimulates it to *Play* (dopamine), *Socialise* (oxytocin), and also be alert (arginine vasopressin) (see Figures 11.7 (b and d)). Then, the user starts stroking the robot several times, provoking oxytocin to promote prosocial behaviours like playing or talking. During the positive interactions, the PB grows to situate at excellent levels (above 90 units). During this period, the robot talks and plays with the user several times, engaging the user and promoting a good relationship. The user mainly provides correct answers in the game, increasing the dopamine levels and the robot's desire to continue playing.

It is worth mentioning that the robot's perception system misclassified the user strokes one time, confusing it with a hit (at second 220 in Figure 11.7 (c)). This fact made the arginine vasopressin levels rise, but the excellent relationship with the user avoids relaxing behaviour (meditation) to become active. Besides, when playing with the robot, the user incorrectly answered one of the robot's questions (around the second 500), causing a notable increase in arginine vasopressin levels. However, the same effect occurs. The PB between both agents is so good that any outstanding negative process becomes enough aroused to become active.

Around 700 seconds after the start of the experiment, the user decides it is time to leave (see Figure 11.7 (c)). At that moment, the levels of the hormones drop back to basal levels (in Figure 11.7 (b)), leading the motivational states related to social interactions to present low intensities. The user's absence leads the robot to not interact with anyone, so it decides to stay calm and wait. However, as Figure 11.7 (a) represents, the PB with the user decays very slowly with time. This effect represents the deterioration of the relationships between humans with time, but in this experiment, the decay has been speeded up, so it is possible to perceive it during a short-term interaction.

During the experiment, the effect that the PB exerts over the cognitive interpretation of the stimuli is perceived in Figure 11.7 (b), where the neuroendocrine response is represented. As the PB rises, the oxytocin response to positive social events is higher, a fact that can be perceived comparing the oxytocin levels at second 100 when the basal value of this hormone is around 0.4 units, and second 700, just before the left of the user, when oxytocin secretion rate is above 0.7 units.

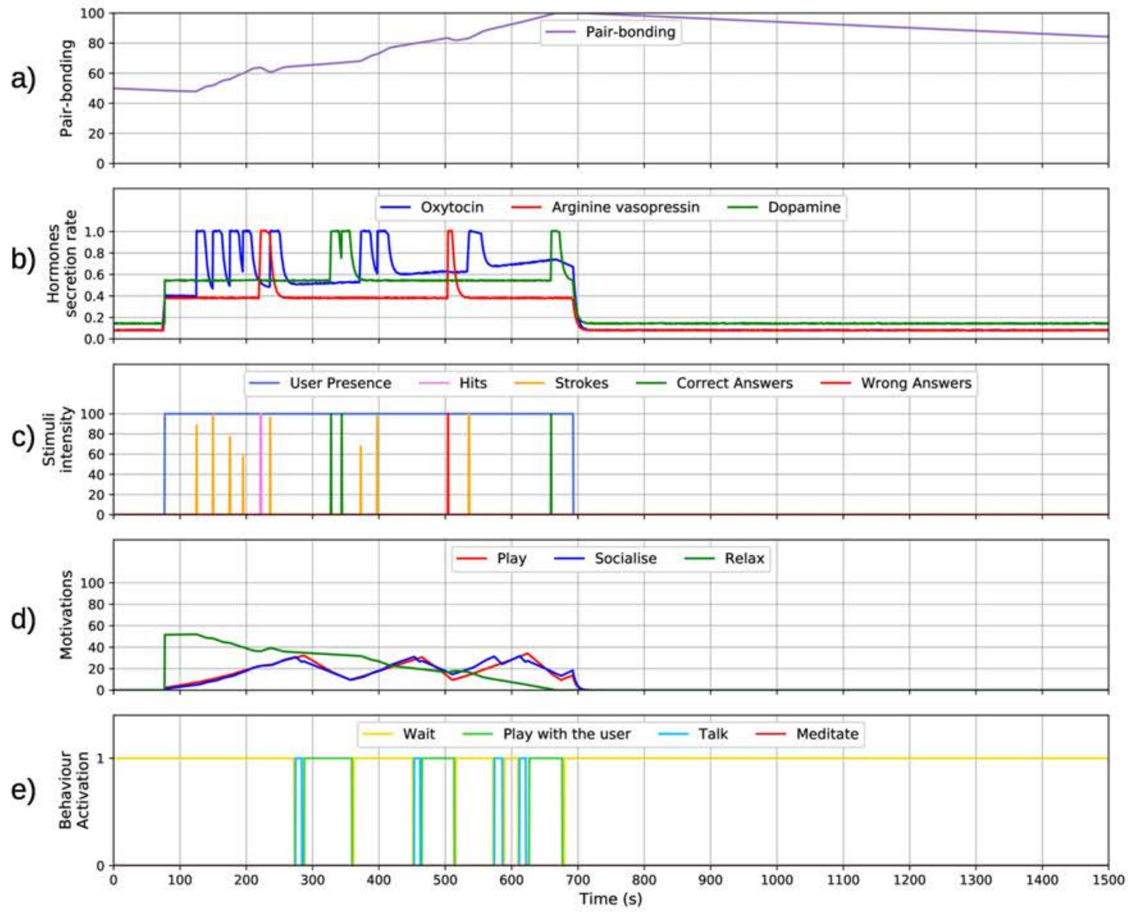


Figure 11.7: Evolution of the PB, hormones, stimuli, motivations, and behaviours of the social robot Mini while interacting with a friendly user.

11.2.3.2 Aversive user

The second round of experiments defined the social robot Mini interacting with an aversive user. It is worth noting that the graphs showing the results where Mini interacted with an aversive user were trimmed to the initial 14 minutes since the most exciting events occurred inside this period. After, the PB was so bad that the robot ignored the user for the rest of the experiment. Figure 11.8 shows the response of the multiple artificial processes to this interaction. As in the previous experiment, the PB value starts as 50 units. Initially, the user is not perceived, appearing in the scene around 50 seconds before the start of the experiment. Right before the user's appearance (represented in Figure 11.8 (c)), the basal levels of the hormones maintain stable around 0.4-0.5 units (dopamine and oxytocin are slightly above arginine vasopressin to foster the interaction). However, the user rapidly starts hitting the robot, provoking a decrease in the hormone involved in social interaction (oxytocin) and the decay of the PB between both agents. Interestingly, the accumulation of continuous hits increases the decay speed of the PB, as they make oxytocin levels drop throughout the impact of the adaptive allostatic mechanism on the cognitive interpretation of the stimuli.

The stress of the robot due to continuous aversive situations translates into a notable increase in the motivation to *Relax* and a substantial drop in the motivations oriented to *Socialise*. This fact drives the robot to ignore the user executing the meditate behaviour. Although the robot ignores the aversive user, its internal state is so bad that it maintains meditating until the user decides to leave (second 480). All hormones go back to their basal values making the robot stop meditating moving to a neutral state.

Around 300 seconds after the start of the experiment, the PB between both agents drops to a null value, not moving above this value in further interactions. The aversive user returns to interact with the robot at second 600 (see Figure 11.8 (c)), but this time, the user interacts positively with the robot. This change in the aversive profile defined in Section 11.2.2 was carried out to investigate whether a seriously deteriorated PB could be restored if a user provides the robot with positive affective stimuli. Nonetheless, providing positive interaction does not cause any benefit on the internal state beyond the stimulation of oxytocin, since a fully deteriorated PB ballasts the emergence of prosocial behaviours like playing or talking. Instead, the robot continues ignoring the user, preferring to meditate. Despite the evolution of the processes shown in this experiment, the robot does not repair the relationship with a user that abused him previously. If the user repeatedly provides positive affect, the relationship can be restored. However, it is not easy in such bad cases as this scenario.

The effect of the adaptive social mechanisms in this scenario is more evident in this case. The most stunning result is the drop in the oxytocin response when the PB becomes terrible (see end of Figure 11.8 (a)). At the beginning of the experiment, the user stroked the robot two times (seconds 95 and 120), raising the oxytocin levels to their maximum value (1 unit per time step). However, right before the finishing of the experiment, the stokes do not provoke the same effect. These times, the stimulation of oxytocin response barely attain secretion rates around 0.5-0.6 units, a fact that is indicative of the influence of the PB on the cognitive interpretation of positive tactile contact and its influence on the oxytocin levels.

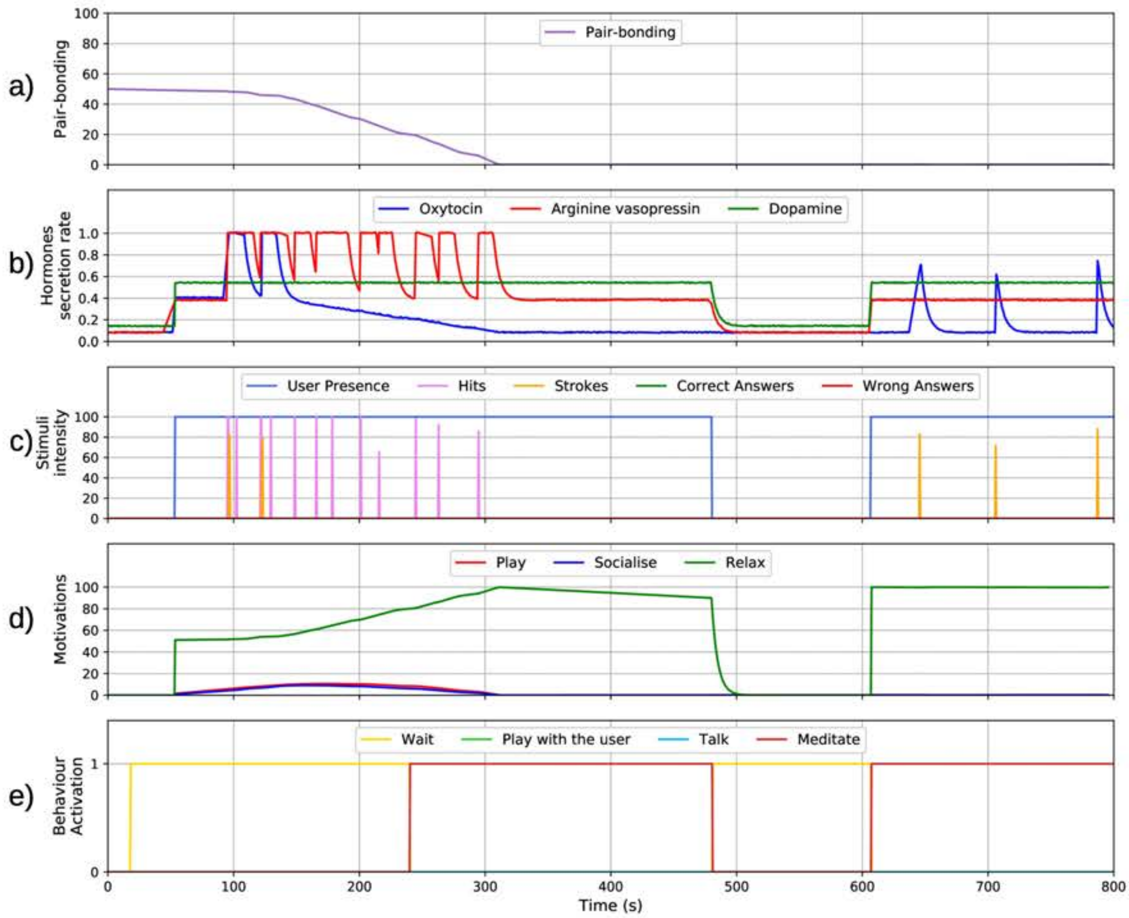


Figure 11.8: Evolution of the main processes involved in the social response of the social robot Mini during the interaction with an aversive user.

11.2.3.3 Neutral user

The last case shows a user who combines the friendly and aversive profiles. Thus, the user sometimes provides positive affective stimuli in the neutral profile and sometimes mistreats the robot. Contrary to the previous two cases, this scenario depicts more general situations rather than presenting extreme-opposite situations. Figure 11.9 shows the evolution of the PB and their related biological processes. As the previous experiments, the user is initially absent, and PB values start at 50 units (Figure 11.9 (a)). Then, the user comes to the scene nearly 80 seconds after the start of the experiment, as shown in Figure 11.9 (c). During the interaction, the user combines hits and strokes at will. Besides, the answers provided when playing together are often incorrect and correct on some occasions. The intersection of both positive and negative stimuli cause the PB to oscillate around moderate levels of bonding. When positive stimuli overcome negative ones, oxytocin and dopamine tend to increase (see Figure 11.9 (b and c)), promoting an increase in the PB and eliciting prosocial behaviours like playing or talking to become active. Since the PB is never too bad, the robot never meditates.

Unlike in the previous cases, a neutral user makes the robot's behaviour more diverse than in previous approaches. Despite the motivation to relax presenting a moderate intensity, it does not become dominant because its activation threshold was set to 80 units to avoid the robot continuously meditating.

In this setting, the allostatic social mechanism does not show determinant outcomes in the evolution of the hormonal responses of the robot. This issue can be explained using the definition of the cognitive modulation parameter β throughout the value of the PB. As stated above, the cognitive modulation parameter β limits or amplifies the influence of a stimulus on the biological process of the robot, presenting a value from 0.01 to 2 units. For this reason, if the PB is around 50 units, the mapping to the β parameter is around 1 units, meaning that the PB does not have an impact on the interpretation of the stimuli. Considering this fact, the evolution followed by the hormones involved in social behaviour raise their levels like the physical modulation parameter α defined in Table 6.2 defines, not presenting a substantial adaptation of its value.

Figure 11.9: Representation of the bonding and related biological process during the interaction of a neutral user and the social robot Mini.

11.2.4. Conclusion

The previous experiments have presented how the social behaviour of a social robot can be adapted to the way the user is behaving with the robot. The adaptation takes inspiration from the allostatic social mechanism described by Schulkin [71] and our initial experiments described in Chapter 8 drawing on Cañamero's research. Using this method, the robot interpretations about the social stimuli change depending on the level of PB between both agents without including learning models. Thus, if the user behaves appropriately with the robot, prosocial behaviours will be elicited more easily, improving the relationship between the user and the robot's internal state. Otherwise, an aversive user deteriorates the social neuroendocrine responses of the robot, driving it to avoid these uncomfortable situations. In the model, Mini avoids aversive situations by meditating, ignoring all the actions carried out by the user.

The results show that the robot can detect whether the user behaves correctly and adapt its behaviour in consequence. On the one hand, oxytocin levels react to positive socialisation, promoting playing with the user and talking with him. On the other hand, arginine vasopressin adapts its secretion rate regulating the robot's response to aversive situations, allowing it to relax when it feels overwhelmed by the negative actions of the user. From our point of view, endowing Mini with these adaptive capabilities may enhance their actuation capabilities in many different ways. First, the robot exhibits a more diverse and natural behaviour that is not predefined, leading people to perceive the robot as a more capable and intelligent machine. Secondly, we believe that the user's engagement might increase, promoting more prolonged interactions and avoiding fatigue. However, these hypotheses must be assessed in real HRI studies, where the robot interacts with a broad group of users whose impressions have to be retrieved using controlled and standardised methods.

The positive outcomes of the adaptive robot behaviour open many chances for studying many other processes involved in the social communication between humans and robots. An example could be the analysis of how the robot's trust may vary depending on the user it is facing, influencing whether the robot follows the user's instructions or concentrates on their own needs and concerns. Considering these new opportunities in modelling the social behaviour of the robot using the allostatic mechanism, as stated in the future work section, the idea is to continue extending the adaptive social capabilities of our robot by endowing them with new biological processes that lead to a more variate behaviour.

CHAPTER 12

Evaluating the robot's expressiveness

This chapter describes the methods and experiments performed for translating the evolution of the biological processes such as emotions, mood, or autonomic processes (e.g. pupil size or breathing rate) into physical responses and expressions to improve the naturalness and liveliness of the robot. The *Motivational model* emulates these biological functions and send their values to the *Expression manager* for its expression using Mini's actuators. The expression of the artificial biological process focuses on two research lines:

- Generating and expressing affective states using emotion and mood.
- Generating and expressing autonomic physiological processes such as the heartbeat using specific actuators of the robot

12.1. Generating and expressing affective states

In the last years, the development of the actuation capabilities of robotic platforms has opened a wide range of possibilities to continue improving their expressiveness. Expressing the robot's needs, intentions, or desires in social robots becomes essential for natural interaction with the human. Previously in this manuscript, we have focused on describing and presenting results about endowing artificial systems with biological capabilities that affect decision-making. Now, we focus on how Mini expresses affect and autonomic processes. In a preliminary review about how to express affect and autonomic processes in social robots, we noticed a lack of research about representing complex affective states (combination of emotion and mood). Besides, we have not found any work attempting to improve the robots' expressiveness using autonomic processes such as heart rate or blinking rate.

Taking advantage of this niche and the emulation of biological processes following a neuroscientific perspective we presented in Chapter 6, we investigated how to express these biological functions using the robot's actuators. Specifically, this chapter addresses the expression of affective states and autonomic physiological processes and their evaluation.

12.1.1. Generating affective states

This dissertation described how the robot's internal state could generate affective states mainly due to external stimuli. In the model, affective states divide into emotions and mood.

12.1.1.1 Emotion

Emotions are reactive psychological states hugely dependent on arousing stimuli, whose effect rapidly disappears once the associated stimulus is no longer perceived. In our model, we use the Lövheim's approach [191], as explained in Section 6.8, to map the levels of dopamine, serotonin, and brain norepinephrine into the emotions *Joy*, *Surprise*, *Sadness*, and *Anger*. As Lövheim posits, these neuroendocrine substances raise their levels with the perception of specific stimuli triggering certain emotions in the robot. As Table 12.1 shows, the four emotions generated by our model have an intensity level from 0 to 100 units that represent how strong the affective response is. Besides, for becoming active, a threshold value of 20 units is used to avoid shallow emotions dominating Mini's affective state.

Emotion	Range	Threshold	Intensity calculation
Joy	0-100	20	$100 * \text{DA}^2 * \text{SE}^2 * (1 - \text{BNE})^2$
Sadness	0-100	20	$100 * (1 - \text{DA})^2 * (1 - \text{SE})^2 * (1 - \text{BNE})^2$
Anger	0-100	20	$100 * \text{DA}^2 * \text{BNE}^2 * (1 - \text{SE})^2$
Surprise	0-100	20	$100 * \text{BNE}^2 * \text{SE}^2 * (1 - \text{DA})^2$

Table 12.1: Emotions generated by our *Motivational model*. DA: Dopamine, BNE: Brain norepinephrine, SE: Serotonin.

12.1.1.2 Mood

Unlike emotion, the mood is a long-term affective state influenced by the accumulation of past experiences lived by the agent. Currently, the affective component of the *Motivational model* can generate nine possibilities using a bidimensional valence-arousal space: *Neutral*, *Happy*, *Excited*, *Angry*, *Afraid*, *Sad*, *Depressed*, *Calmed*, and *Contented* [195] (see Figure 12.1). The *Motivational model* uses two biological processes called pleasantness and arousal to define the value of the valence and arousal axes that define Mini's mood, as explained in Section 6.8. On the one hand, the robot's pleasantness, that is calculated using Equation 12.1, depends on how bad the robot's deficits (d_i) are:

$$\text{pleasantness} = \sum_{i=0}^{N} d_i \tag{12.1}$$

where N is the number of biological processes with a deficit associated.

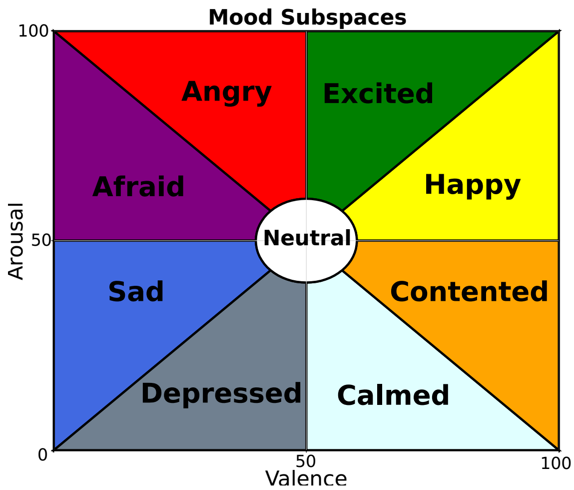


Figure 12.1: Moods in the valence-arousal space.

On the other hand, Equation 12.2 expresses how the arousal value is calculated. In our approach, we propose that the robot's arousal depends on its Stress and Wakefulness levels, two variables implicated in the robot's excitement.

$$\text{arousal} = \frac{\text{Stress} + \text{Wakefulness}}{2} \tag{12.2}$$

12.1.1.3 Mini's affective state

The emotion and mood generation derives from Mini's affective state (*Affect*). In this thesis, we propose that Mini's affective state combines its dominant emotional state and its current mood, as Equation 12.3 shows. On the one hand, the dominant emotional state is the emotion with the highest level of intensity among all active emotions (those with an intensity level about the threshold of 20 units). On the other hand, the current mood depends on the pleasantness and arousal values in the valence-arousal bidimensional space. The Motivational model sends the robot's affective state to the *Expression manager* (see Section 5.2 for further information), a module in the robot's software architecture in charge of expressing affective states using the robot's actuators.

$$\text{Affect} = \begin{cases} \text{Dominant emotion} \\ \text{Current mood} \end{cases} \tag{12.3}$$

12.1.2. Experimental setup and evaluation

The experiment for evaluating the affective states expressiveness of the social robot Mini was performed to assess:

- The emotion and mood generation (affective state) in Mini's from its internal processes and the stimuli the robot perceives.

- If Mini's affective state was coherent with the situation the robot was experiencing.
- If people correctly perceive the affective states that Mini is expressing.

In this thesis, we focus on emotion and mood generation since the robot expressions is part of another research line [300] not related to this thesis.

As we explained in the previous section, the emotion and mood generation defines the robot's affective state, and the *Expression manager* controls the robot's actuators to express what Mini is feeling. We designed predefined expressions to each emotion and mood for expressing affective states. These expressions are predefined profiles that indicate how each robot actuator should operate depending on the current robot's affective state. Since we designed our affective expressions and adjusted them following our impressions, we designed an experiment to evaluate them with people. The experiments followed two stages:

- **Experiment 1:** The first stage consisted of spreading an online survey containing videos showing the robot expressing the previously mentioned emotions and mood. In total, 10 videos were displayed showing:
  - Punctual emotional reactions like *Anger*, *Joy*, *Sadness*, and *Surprise* (4 videos).
  - The decay of the previous emotions with time (4 videos).
  - The robot expressing a Neutral and a Happy mood (2 videos).

  To facilitate affective expressiveness, we mapped the perception of each arousing stimuli to a specific emotion. Each video had a duration between 20 to 30 seconds. Then, the participants had to select which emotions or moods they perceived in the robot using a predefined list of options. In the case of the emotions, the alternatives shown to the user were the six basic emotions of Ekman [185]. In the case of the moods, the predefined options were *Happy*, *Neutral*, bored, anxious, and relaxed, a list obtained from the Mehrabian approach presented in [195]. This stage allowed us to check if the participants could correctly perceive the affective states of the robot individually since in the **Experiment 2** we designed a more complex scenario using HRI. In this preliminary study, 55 participants completed the online survey. Then, using the survey's suggestions, we improved those aspects of the affective expressiveness that the participants felt as inconsistent (e.g. the voice of the robot).

- **Experiment 2:** Next, we carried out a more thorough experiment to assess different ways of expressing affect during HRI. The experiment consisted of four videos of around 5 minutes of duration where the robot's affective state was evaluated using different approaches. All videos showed an experimenter playing with Mini a quiz game. First, the robot introduced the experimenter to the activity they would execute. Then, the experimenter had to answer four questions about historical facts using a touch screen during the game. Each question had a correct answer from a predefined list of four alternatives. During the formulation of the second and the fourth questions, the experimenter hit and stroked the robot, provoking the robot's reaction, which could be affective in some of the conditions tested. The four conditions assessed were the following:

- A *Neutral* robot not showing any kind of affective states (Condition 1: *Neutral*).
- A robot only reacting showing punctual emotions as a response to the different stimuli (Condition 2: emotions).
- A robot showing punctual emotional responses and mood states (Condition 3: emotions and mood.
- A robot showing emotional punctual responses with decay and mood states (Condition 4: emotional decay and mood).

Hypothetically, the Condition 4 should be the one showing the best results since it combines the expression of emotional decay and mood states. Due to the Covid-19 pandemic, the evaluation was performed using an online survey between users. After watching one of the four videos showing one of the conditions previously described, the participants of the study had to complete the well-known RoSAS [359] standard questionnaire (included in Appendix D.3), a tool used for evaluating different categories of social robots. The study participants were randomly assigned to one of the four described conditions. In this experiment, 30 people completed the survey assigned to the first condition, 34 to the second one, 35 to the third condition, and 33 to the fourth condition.

The evaluation of the **Experiment 1** was carried out to appraise if the individual expressions of Mini were correct was performed using an online survey. Although the Motivation model can generate four emotions (*Joy*, *Surprise*, *Sadness*, and *Anger*) and nine moods (*Neutral*, *Happy*, *Excited*, *Angry*, *Afraid*, *Sad*, *Depressed*, *Calmed*, and *Contented*), we just evaluated two moods (*Neutral* and *Happy*) as the *Expression manager* does not include expressions for the other moods yet.

The **Experiment 2** was conducted to assess whether a more detailed expressiveness of the social robot Mini improves the study participants' perception of the robot during a long-term interaction. The RoSAS [359] questionnaire measures the perception of the robot's users using multiple questions classified in three categories: warmth, competence and discomfort. Since there is no official Spanish version of the RoSAS questionnaire (in Appendix D.3), we translated it ourselves to conduct the study. Our initial hypothesis was that the robot showing emotional decay and mood (Condition 4) should provide the best rating in warmth and competence and the lower ones in discomfort. Besides, we expected that the comparison between the four conditions show significant differences in the warmth and discomfort categories, but not in the competence one since the robot presented the same capabilities in all the conditions tested. Thus, the comparison of the four conditions was statistically analysed to investigate whether endowing the social robot Mini with a more natural and complete affective expressiveness let the participants perceive the robot as warmer and less comfortable.

Figure 10.7 shows the evolution of the robot's affective state in the video shown to each participant of the study. Since four conditions were evaluated, four videos were shown with subtle differences among them. The videos showed the experimenter interacting with Mini while playing a quiz game. After a short introduction about the game, the robot asked four history questions that the experimenter had to answer. After each answer, the robot explained the historical fact of the question. After question 3, the experimenter hit the robot because it was disappointed by providing a wrong answer. After question 4, the experimenter stroked the robot after guessing the answer. The scenario was designed for the robot to show the four emotion presented above after each question (*Joy*, *Sadness*, *Anger*, and *Surprise*). Besides, depending on the condition evaluated, the robot in the video expressed mood and emotional decay.

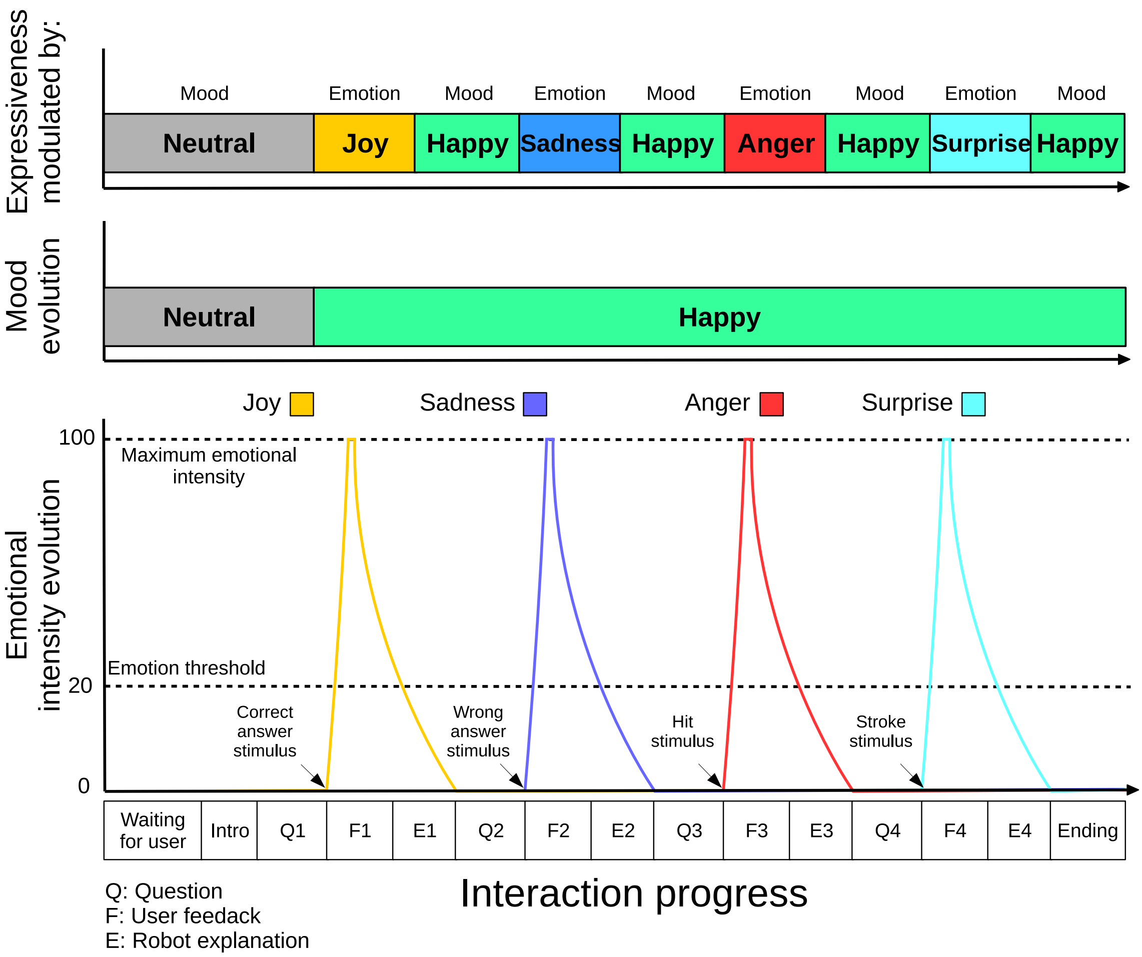


Figure 12.2: Mini's affective state evolution during the long-term interaction.

12.1.3. Results

Following, we present the results of the two evaluations we carried out for the generation and expression of affective states:

- **Experiment 1:** a preliminary study to evaluate individual affective expressions.
- **Experiment 2:** HRI study to evaluate affective expressiveness during a long-term interaction.

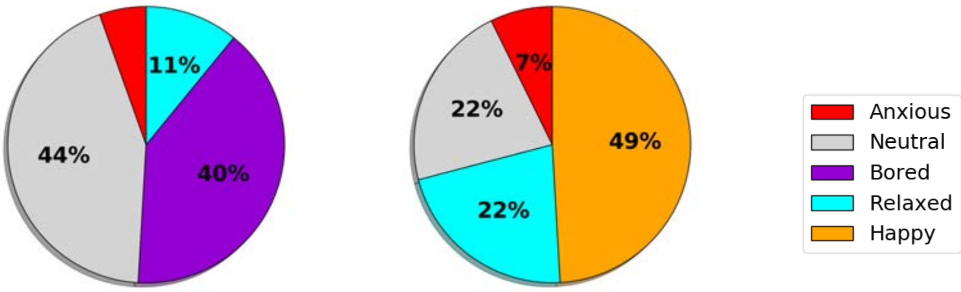
12.1.3.1 Experiment 1: preliminary study

As described above, in the first experiment we evaluated three different cases:

- The robot expressing a *Neutral* and a *Happy* mood.
- Punctual emotional reactions like *Anger*, *Joy*, *Sadness*, and *Surprise*.
- The decay of the previous emotions with time.

Next, we show the results obtained in each evaluation.

- The preliminary study provided very positive results in the recognition of the mood displayed by the robot (see Figure 12.3). More specifically, 44% of the participants correctly perceived the *Neutral* mood, despite 40% of the participants confused it with a *Boring* mood. These results can be explained considering the participants' expectancies from the robot. Since the robot does not express any affective cue in a *Neutral* mood, some participants may perceive it as boring, as both moods present very similar features on their activation level. On the other hand, the participants found easier the recognition of the *Happy* mood. Results show how 49% of them successfully perceived a *Happy* robot, followed by 22% recognising *Neutral* and *Relaxed* mood states. These initial results in terms of mood expressiveness validated the individual expressions regarding the two mood states that Mini can show.



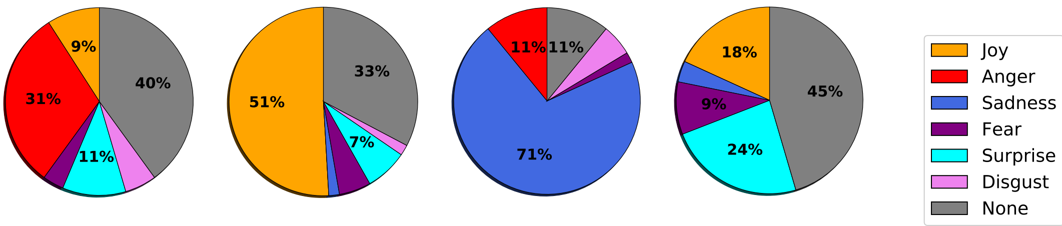
(a) Neutral. (b) Happiness.

Figure 12.3: Results regarding the recognition of the mood states of the robot by the participants of the study.

- The results about the robot showing emotions and their decay with time are shown in Figure 12.4. Participants had several difficulties recognising the robot's emotions in most of the cases. In the video where the robot expressed *Anger*, the emotion was correctly perceived by 31% of the participants. Nevertheless, participants selected *Surprise* with the 11%, *Joy* 9%, disgust 6%, and *Fear* 3%. 40% of participants indicated that they did not perceive any emotion in the robot. Joy yielded better results since 51% of participants correctly indicated the emotion the robot was expressing.

  Nonetheless, 33% of participants did not identify any emotional state, 7% identified *Surprise*, 5% *Fear*, and 2% *Sadness* and disgust. Sadness was the emotion more easily recognised by the study participants with a success rate of 71%. The rest of the users perceived *Anger* 11%, 4% disgust, and 3% *Fear*. 11% of users did not perceive any specific emotional state in the robot during the visualisation of the video. The study participants have significant problems identifying *Surprise* in the robot since just 24% correctly recognised the emotion.

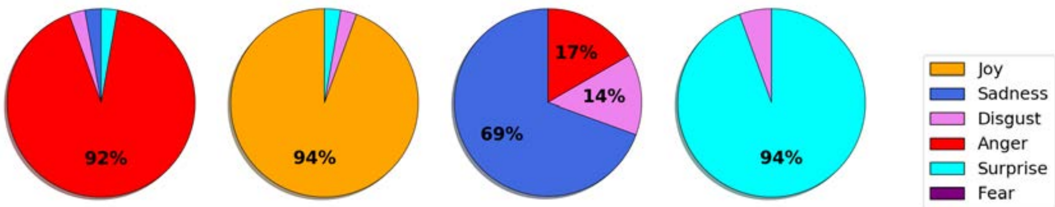
  Moreover, 45% of people did not perceive any emotional state in the robot, 18% perceived *Joy*, 9% *Fear*, and 4% *Sadness*. The lack of context during the evaluation probably led people not to perceive any emotion in some situations. Nonetheless, results show that, for all emotions, the correct answer was always selected as the first (*Joy* and *Sadness*) or second (*Anger* and *Surprise*) alternative. Besides, including emotional decay makes the emotional intensity reduce its effect on the robot's expressiveness, maybe letting people think that the robot is not expressing any emotion at the end of the affective event.



(a) Anger. (b) Joy. (c) Sadness. (d) Surprise.

Figure 12.4: Perceptions of the participants in the evaluation of Mini expressing emotional decay with time –*Anger*, *Joy*, *Sadness*, and *Surprise*– in a multi-choice close question approach.

- Focusing now on the punctual emotional expressions, the results show an excellent percentage of successful recognition in *Anger*, *Joy*, and *Surprise*. However, the participants found the recognition of *Sadness* more difficult, as depicted in Figure 12.5. Participants recognise *Anger* in the 92% of the evaluations, *Joy* in 94% of the situations, and *Surprise* in 94% of the cases. In all these evaluations, the other emotions had residual percentages in the recognition. The emotion of *Sadness* was recognised by 69% of the participants, followed by *Anger* (17%) and *Disgust* (14%). Despite recognising *Sadness* provided results below the other emotional states, its identification percentage is still very high. Unfortunately, the participants did not provide any clue about why they found the recognition of *Sadness* more complex than the others.



(a) Anger. (b) Joy. (c) Sadness. (d) Surprise.

Figure 12.5: Values obtained in the evaluation of the emotional expression of the robot -*Anger*, *Joy*, *Sadness*, and *Surprise*- in a multi-choice close question approach.

The comments section included in the questionnaires used for the previous evaluation suggested a couple of points that we addressed for the second experiment:

- Many participants perceived that the voice modulation of the robot was not adequate to the affective state the robot was expressing. In this case, we change some parameters of the voice modulation following the participants' instructions.

- Some participants reported the importance of the eyes' expressiveness to correctly perceive the robot's affective states.

12.1.3.2 Experiment 2: HRI study

Figure 12.6 shows the results of the evaluation of the robot's affective expressiveness in a HRI scenario considering 4 different conditions and the RoSAS questionnaire [359].

- Condition 1 shows a robot not expressing any affective expressiveness.
- Condition 2 shows a robot with emotional responses.
- Condition 3 a robot with emotional reactions and mood.
- Condition 4 a robot with emotional decay and mood.

Since the data distribution did not follow a normal distribution, we performed non-parametric analysis using the Kruskal-Wallis [360] test for several independent samples. Unexpectedly, like Table 12.2 shows, no significant differences were obtained in the comparison, so our initial hypothesis about finding differences in the warmth and discomfort categories of the RoSAS questionnaire was not correct. The Condition 3 that defined a robot showing emotions and mood but without emotional decay outperformed Condition 4, where the robot affective expressiveness included emotional decay. These results refuse our initial hypothesis that stated the benefits of including emotional reactions with decay and mood states in the affective expressiveness of the social robot Mini.

RoSAS Attribute	Kruskal-Wallis H parameter	DoF	p-value
Warmth	4,056	6	0,255
Competence	3,958	6	0,266
Discomfort	4,338	6	0,227

Table 12.2: Kruskal-Wallis tests for the RoSAS attributes warmth, competence, and discomfort regarding the comparison of the four conditions tested. C1: Robot without affective regulation, C2: Robot with emotional reactions to stimuli, C3: Robot with emotional reactions and mood regulation, C4: Robot with emotional reactions decaying with time and mood regulation.

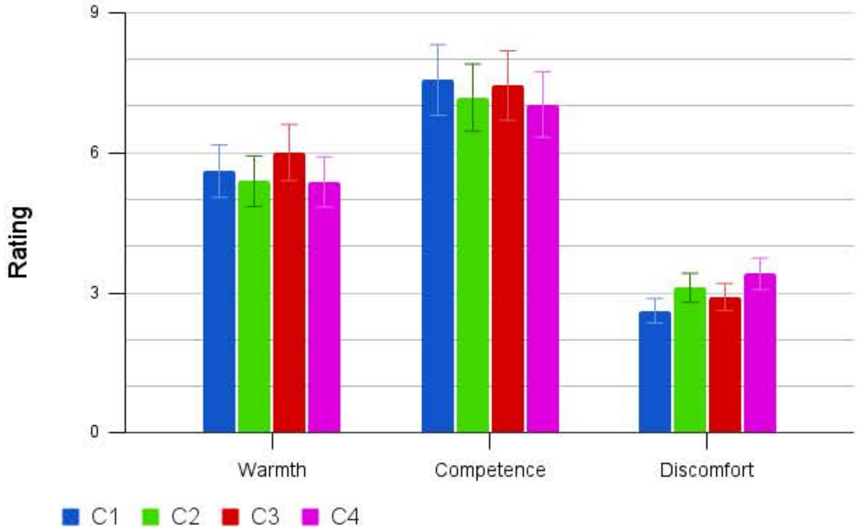


Figure 12.6: Statistical analysis (mean-standard deviation) of the four conditions tested in the affective expressiveness experiment. C1: Robot without affective regulation, C2: Robot with emotional reactions to stimuli, C3: Robot with emotional reactions and mood regulation, C4: Robot with emotional reactions decaying with time and mood regulation.

12.1.4. Discussion

Endowing social robots with affective responses has been deeply studied in social robotics. However, most of the current literature in affective expressiveness focus on emotional reactions, leaving aside the temporal evolution that affective states undergo in humans. For this reason, the previous sections presented two experiments for assessing the opinion of the robot users about the affective expressiveness of the robot. It is worthy to note that our contribution to this work resides in the affective generation of emotional and mood states, leaving the designing of the affective expressions in the hand of another research line more focused on the expressiveness of social robots. Next, we discuss the most interesting outcomes of the two experiments.

- The results of the **Experiment 1** provided valuable outcomes. The study participants could satisfactorily identify the mood and emotions as a reaction to the social robot Mini's stimuli. However, the results showed that emotions experiencing a temporal decay unrelated to specific stimuli are difficult to perceive. The recognition of the robot's mood was successful in almost all attempts, but participants had more difficulties than recognising punctual emotional reactions due to the decay of emotions with time. Nevertheless, recognising a *Happy* mood yielded better results than the *Neutral* one, probably because the *Neutral* mood can be easily confused with similar mood states like calmed or bored. Secondly, recognising the emotional responses of *Anger*, *Joy*, and *Surprise* yielded excellent results since their recognition percentages were all above 90% per cent of success.

Although the recognition of *Sadness* did not provide such excellent results, its recognition was acceptable, attaining a recognition percentage slightly below 70%.

Apart from the evaluation of the individual affective expressions of the robot, this preliminary study allowed us to obtain the participants' suggestions about the robot's expressiveness. Most participants suggested that the robot's voice was not in synchrony with the affective expression that the robot was transmitting. Besides, a couple of participants suggested that the expressions of the eyes are important for recognising the robot's affective state since it provides meaningful information. Considering these suggestions, the affective expressions of the robot were improved, focusing on the voice and eyes expressiveness.

- The second experiment we conducted was to test how a new group of participants perceived the expressiveness of the robot's affective state using four different approaches. Each approach increased the level of detail and complexity of the affective expressiveness, moving from a robot without this capacity to an affective generation system that could express emotional reactions, the decay with time of such emotions, and the mood of the robot in a real HRI case. However, the results of this second study did not reach our expectations. The evaluation did not show any significant statistical differences in comparing the four approaches in any of the categories evaluated: warmth, competence and discomfort. Besides, contrary to our initial hypothesis, the robot's condition showed emotional reactions and mood but not emotional decay was the best rated. The possible causes behind these results may be diverse. In the first place, the scenario we designed for evaluating the affective expressions might not be appropriate since the participants could not interact face to face with the robot. Besides, the limitations of the social robot Mini's actuation capabilities might bound the robot's expressiveness. For this reason, we believe that the evaluation of Mini's affective expressiveness might require further experimentation to fulfil our initial hypothesis. As the expressiveness research was not mainstream, the future work section tackles this issue.

12.2. Generating and expressing autonomic processes

This section describes the generation and expressiveness of autonomic processes in the social robot Mini. We begin by defining how affective states are elicited in the robot. Then, we show the experiment and evaluation of the evolution of these processes in Mini. Finally, we show the results about the evolution and expressiveness of these processes during long periods.

12.2.1. Generating autonomic processes

The review of the existing literature in modelling biological processes in embodied systems shows an evident lack in translating autonomic physiological processes into accurate representations using the robot's actuation. For this reason, we have focused on studying how to simulate physiological responses and regulate autonomic responses like the heart rate or pupil size in the robot. Although many autonomic physiological processes impact human expressiveness, we are limited by the actuation capabilities of Mini in this work. Hence, we concentrated on modelling and representing the evolution of its heart rate, pupil size, blinking rate, breathing rate, and locomotor activity. As Section 2.3.1 described and Table 12.3 shows, the evolution of these processes depends upon circadian and ultradian rhythms, the intensity of specific stimuli, and the secretion of neuroendocrine substances.

Autonomic process	Range	Circadian rhythm	Activation	References
Heart rate	[50, 200] beats per minute	$60 + 3 \cdot \cos(2\pi \cdot (\frac{t-12}{24}))$	$100 * \frac{(EPI+ANE)}{2}$	[144, 313] [314]
Breathing rate	[10, 30] breaths per minute	$20 + 5 \cdot \cos(2\pi \cdot (\frac{t-7}{24}))$	$5 * CRH - 5 * DA$	[315] [316] [317]
Blink rate	[0, 40] blinks per minute	$20 + 10 \cdot \cos(2\pi \cdot (\frac{t-22}{24}))$	$10 * DA * SE$	[318] [319]
Locomotor activation	[0, 100] units	$50 + 1 \cdot \cos(2\pi \cdot (\frac{t-7}{24}))$	$100 * \frac{(DA+SE)}{2}$	[144] [320]
Pupil size	[2, 8] mm	None	$ANE + 3 * EPI$	[314, 321]

Table 12.3: Autonomic processes included in the social robot Mini. EPI: Epinephrine, ANE: Adrenal norepinephrine, CRH: Corticotropin realising-hormone, DA: Dopamine, SE: Serotonin.

Drawing on relevant studies in neuroscience previously presented in Section 6.5, we show the evolution of autonomic processes as reactive variables with a natural day period. The evaluation of the expressiveness of the affective states and the autonomic physiological process was assessed in two independent experiments. The following section provides a detailed description of both experiments.

12.2.2. Experimental setup

The control of the autonomic expressiveness of the robot was evaluated using a proof of concept to analyse the evolution of the autonomic processes and their expression in the robot. The graphs showing the evolution of the autonomic processes considered the daily evolution ($24h$) of the circadian pattern of the process, the effect of particular chemicals in sudden increases and decays of the processes, and the influence of stimuli on their autonomic regulation. Besides, we provide several figures depicting the social robot Mini displaying the autonomic responses. As mentioned before, we concentrated on showing the regulation of:

- **Heart rate:** In Mini, the heartbeat controls the fading up and down of a LED that emulates a coloured heart.
- **Pupil size:** The pupil size modulates the diameter of the artificial pupil of the robot.
- **Blinking frequency:** The blinking frequency regulates the frequency with which the robot blinks.
- **Breathing rate:** The breathing rate autonomic process controls the frequency and volume of a sound, simulating the respiration of the robot.
- **Locomotor activation:** The locomotor activity affects the frequency and amplitude of the robots' movements, representing the robot's arousal.

12.2.3. Evaluation

The evaluation focused on representing the evolution of these processes on the robot and in a series of graphs that allow us to see at a glance the causes that make autonomic processes evolve in a certain way during a natural day according to the mathematical equations defined in Section 6.5 that consider many internal and external simulated biological process to yield autonomic reactions. In all the scenarios we evaluate, the graphs show:

a) The evolution of the autonomic process.

b) The levels of the hormones that regulate the autonomic process.

c) The stimuli with influence on the autonomic process.

Finally, we present a general case showing the evolution of all autonomic processes running in Mini. This final representation allows us to evaluate the autonomic control in full detail to compare the automatic reactions of the robot.

12.2.4. Results

The emulation of autonomic processes in the social robot Mini is translated into accurate physiological signals that simulate specific expressions of the social robot. The following paragraphs show the results in regulating the five autonomic processes in the social robot Mini: heart rate, pupil size, blinking frequency, breathing rate, and locomotor activity.

12.2.4.1 Mini's heart rate

Figure 12.7 represents the evolution of the heart rate signal during 24h. The evolution depends on a fixed component defined by the circadian rhythm of the heart and a reactive component that represents the secretion values of hormones and neurotransmitters. On the one hand, the circadian pattern of the heart rate is mathematically expressed using a cosine function that peaks in the early morning (around 11am) and have a nadir in the early night (around 12pm). The heart rate value is around $55-65$ beats per minute (bpm) in normal conditions (see Figure 12.7 (a)). In humans, heart rate values below 50 bpm and above 160 bpm are considered abnormal, so in our model, these values are the lower and upper bounds of the variable. In our model, epinephrine and norepinephrine are important stimulators of heart rate values. These chemicals are secreted in stressful situations, leading the body to provide an autonomic response to restabilise the body's deficiencies. Considering this fact, the levels of these two substances rise with the perception of arousing stimuli in the environment. In this work, the aversive user and hits that the robot receives are the stimulators of epinephrine and norepinephrine. Unexpected risings of these substances produce a subsequent increase in heart rate, like Figure 12.7 (b) shows. In this evolution, the robot perceives an aversive user at 4pm (Figure 12.7 (c)) that increases the heart rate via the action of norepinephrine and epinephrine (Figure 12.7 (b)). Besides, the aversive user continuously hits the robot, making this increase even more notorious. When the aversive user is negatively interacting with the robot Figure 12.7 (c), the heart rate reaches values above 120 bpm, close to 140 bpm when the user hits the robot with high intensity. Once the aversive user is no longer perceived, the heart rate restores to its basal values.

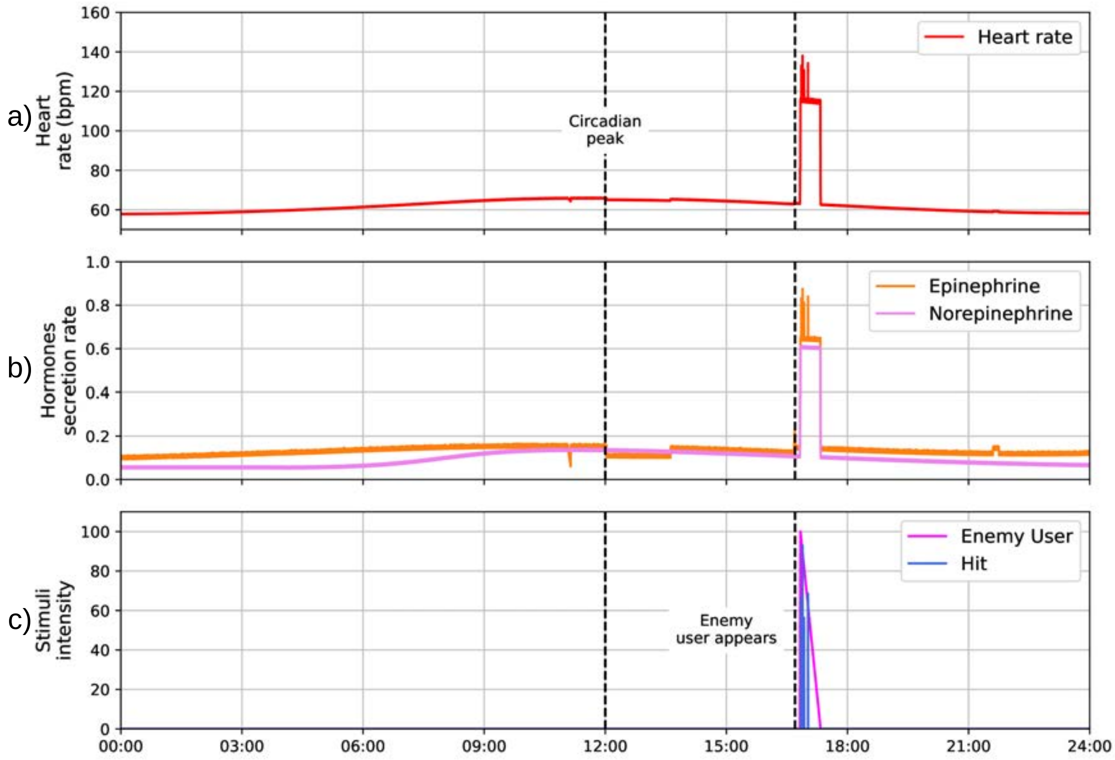


Figure 12.7: Representation of the signals involved in the emulation of the robot's heart rate. On top of that, the heart rate is a daily variation in bpm. In the middle, the secretion rates of epinephrine and norepinephrine during a natural day. Below is the intensity of the stimuli that influence the heart rate.

In the robot, the previous heart rate value is physically represented using a LED that emulates the heart beating. Figure 12.8 shows the fading of the robot's heart with different intensities.

Figure 12.8: Embodied representation of the heart rate in the social robot Mini. The heart rate fades up and down to emulate the heart beating. The frequency of the beating is controlled by the heart rate autonomic process.

12.2.4.2 Response of Mini's pupils

The autonomic control of the social robot Mini pupil size is modelled considering the influence of the light intensity and the substances epinephrine and norepinephrine. Unlike other autonomic processes, pupil size does not present its circadian rhythm, but the light level modulates its daily variation. According to a physiological study, [314], in humans, the pupil size varies from 2 to 8 mm. The pupil dilates in the absence of light and constricts with bright lights. Similarly, the stimulation of the noradrenergic and adrenergic receptors in the spinal cord provoke the dilation of the pupil, normally as a response to stressful situations. Considering these effects, Figure 12.9 shows the daily evolution of the pupil size of the social robot Mini during a natural day. The pupil size presents values around 5 mm during the light period in normal light conditions. Nevertheless, the pupil fully dilates to its maximum 8 mm during the night. Exposing the robot to bright light lead the pupils to constrict to a value that depends on the intensity of the exposure. Figure 12.9 (a) shows this situation at 11pm. In the dark period, sudden exposure to bright light provokes a fast adaptation of the pupil size as a defence mechanism, avoiding the iris's damage. It is worthy of mentioning that when the robot is sleeping, its pupils completely dilate, although it can not be appreciated.

The perception of stressful stimuli makes epinephrine and norepinephrine levels rise, driving the pupil to dilate. Figure 12.9 (b) shows, in our system, the levels of norepinephrine and epinephrine rise when an aversive user appears and hits the robot at around 4pm. However, the perception of a friendly user at 12pm subtly provokes the opposite effects, decreasing the basal secretion of both chemicals leading to a very smooth pupil contraction. According to our modelling, the impact of unexpected traumatic events causes a significant dilation of the pupil size, as we pretend that the users are aware of the robot's feelings.

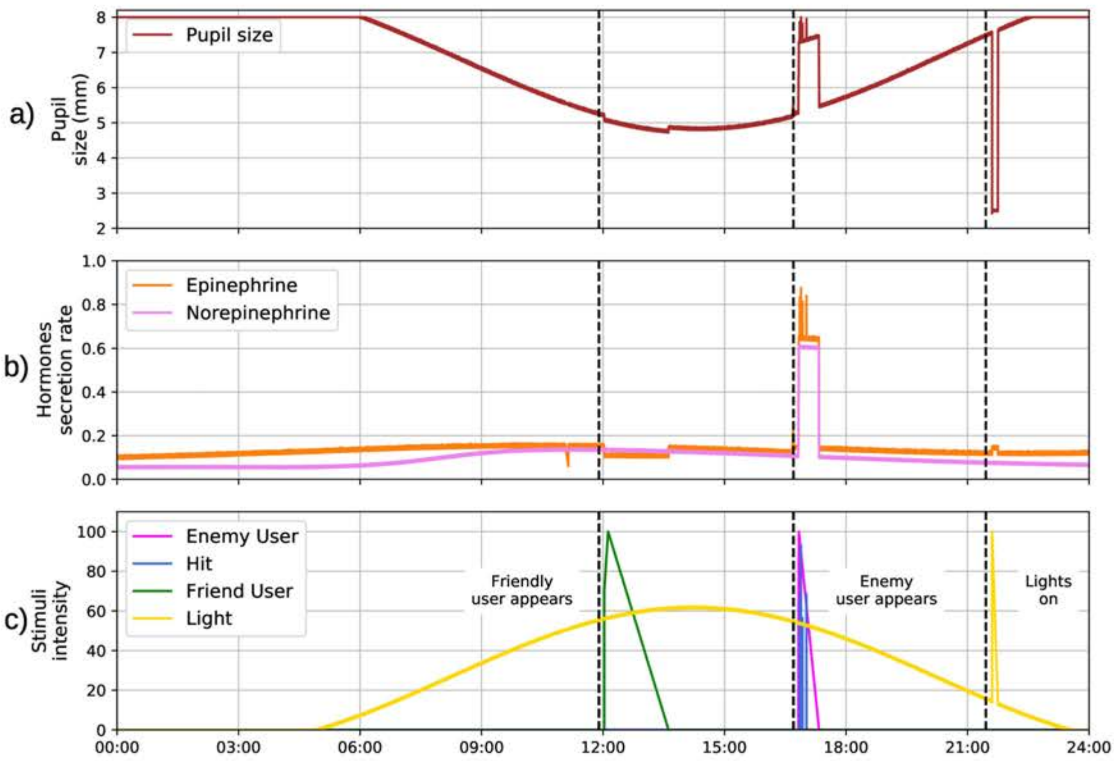
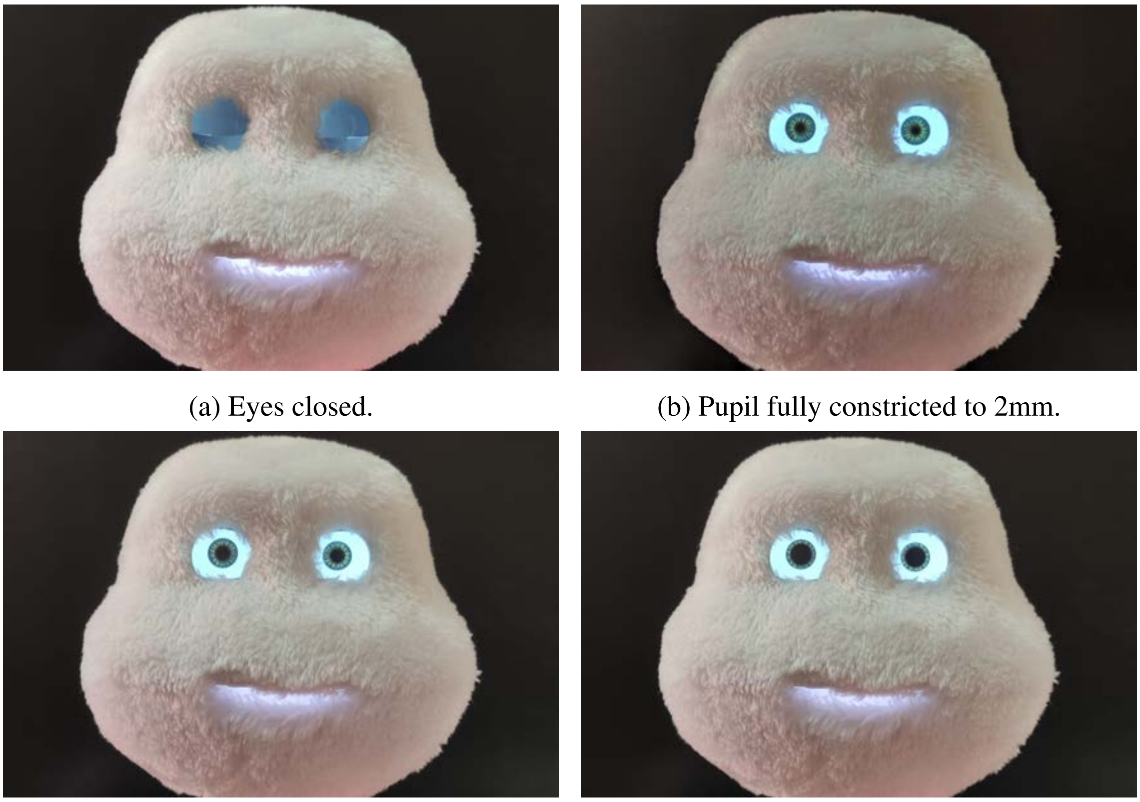


Figure 12.9: Evolution of pupil size during a natural day in the social robot Mini. Top) Reactive size of the pupil due to the light intensity perceived by the robot and the secretion rate of epinephrine and norepinephrine. Middle) Evolution of the secretion rate of epinephrine and epinephrine responds to a threatening situation and circadian variations. Bottom) Illumination level perceived by the robot. Signals represent an enemy user's perception and hit received by the robot.

Figure 12.10 shows the pupil dilation in the social robot Mini. In case the robot sleeps, the eyelids are closed. However, when external stimuli arouse the neuroendocrine system of the robot, the pupil size adapts according to the nature and intensity of the stimuli.



(a) Eyes closed. (b) Pupil fully constricted to 2mm.

(c) Pupil half-dilation to 5mm. (d) Pupil maximally dilated to 8mm.

Figure 12.10: Representation of the changes in the pupil size of the social robot Mini. The size of the pupil depends on the light intensity and the secretion rates of epinephrine and norepinephrine, to important indicators of the perception of aversive stimuli.

12.2.4.3 Mini's blinking rate

Dopamine and serotonin secretion value strongly impact the evolution of the autonomic blink rate. Both substances increase their secretion rates when a positive stimulus is perceived. Besides, serotonin levels also increase when the agent feels negative arousing stimuli. Figure 12.11 (a) shows the evolution of the blink frequency during a natural day. In this graph, a clear circadian pattern modelled using a cosine waveform defines the basal signal that presents a peak before going to sleep and a nadir during the light hours. Like Figure 12.11 shows, the perception of a friendly user and positive touch contact causes an increase in dopamine and serotonin that, at the same time, translates into an increase in the blinking rate.

Similarly, harmful touch contact around 4pm raises serotonin levels, yielding a more subtle increase in the blinking frequency when dopamine levels also peak. Blink rate values range from 10 to 40 blinks per minute. However, typical blinking rates situate around 20 blinks per minute. It is worth mentioning that the eyelids do not blink when the robot is sleeping. However, Figure 12.11 shows an ideal case where the robot does not sleep during a natural day (24h).

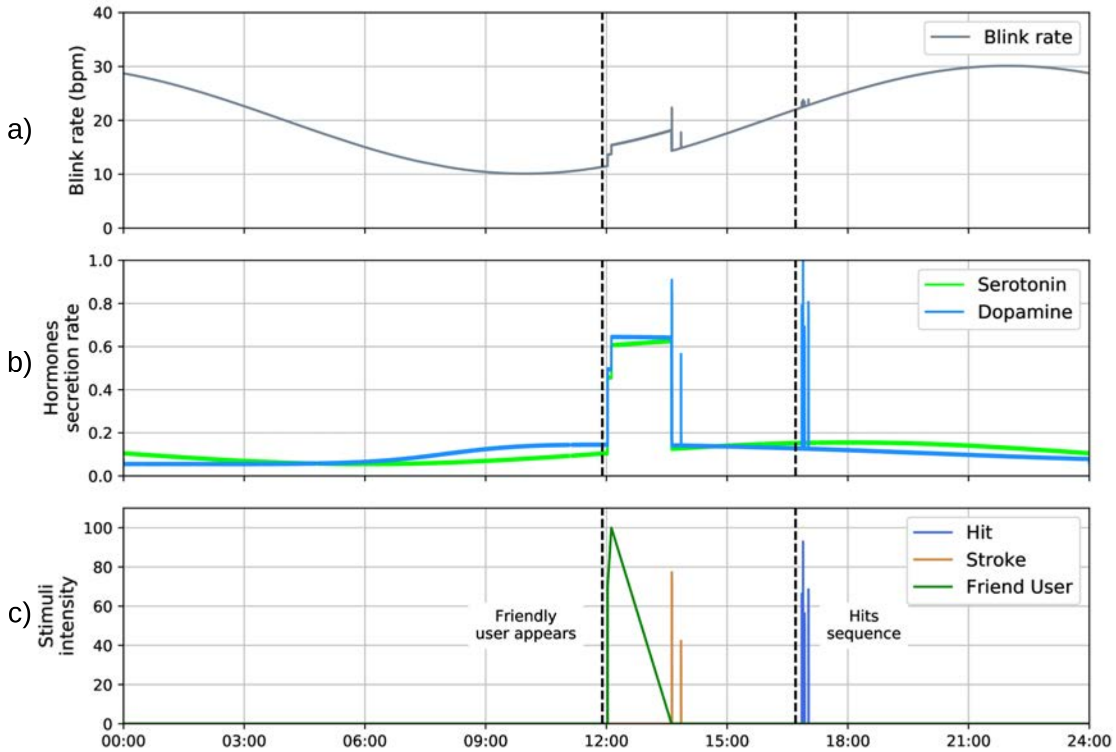


Figure 12.11: Blink rate evolution in the social robot Mini. The neuromodulators dopamine and serotonin regulate the blink rate throughout the indirect action of multiple external stimuli.

Figure 12.12 shows a blink sequence in the social robot Mini. This process repeats at a rate defined by the autonomic blink rate unless the robot sleeps.

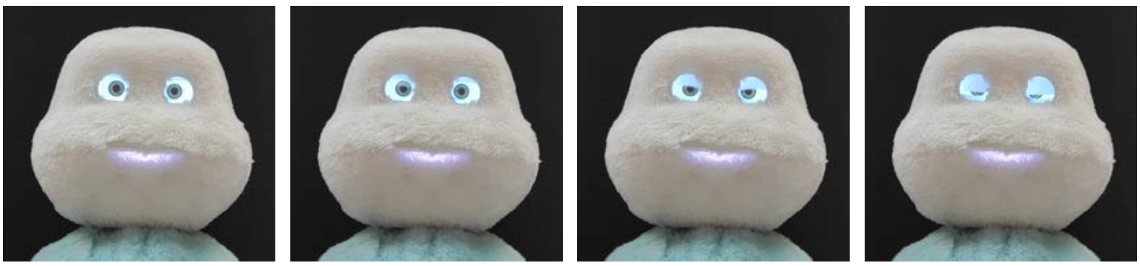


Figure 12.12: Representation of a blinking sequence in the social robot Mini.

12.2.4.4 Mini's breathing rate

The evolution of the breathing frequency and their associated stimulators in the social robot Mini are shown in Figure 12.13. The evolution of the breathing rate is expressed in breathings per minute (bpm). As Figure 12.13 (c) shows, there are many stimuli with a direct implication in the breathing speed of the robot. On the one hand, the presence of a friendly user that strokes the robot stimulates dopamine secretion and inhibits CRH (Figure 12.13 (b)), relaxing the robot provoking a decrease in the number of breathings per minute. On the other hand, an aversive user that hits the robot raises the levels of dopamine and CRH, arousing the robot and, consequently, increasing its breathing rate. Besides, the exposure to bright light stresses the robot, increasing the levels of CRH and the breathing rate. Although the breathing rate is a very reactive process, in normal conditions, it exhibits a precise circadian rhythm that peaks in the early morning (around 7 am) and has lower values in the late evening (around 7 pm) (Figure 12.13 (a)). The breathing rate frequency ranges from 10 to 30 breaths per minute. In Mini, the breathing rate is expressed using a sound that simulates the breathing of the robot. According to the breaths per minute, the sound is modulated in amplitude and duration.

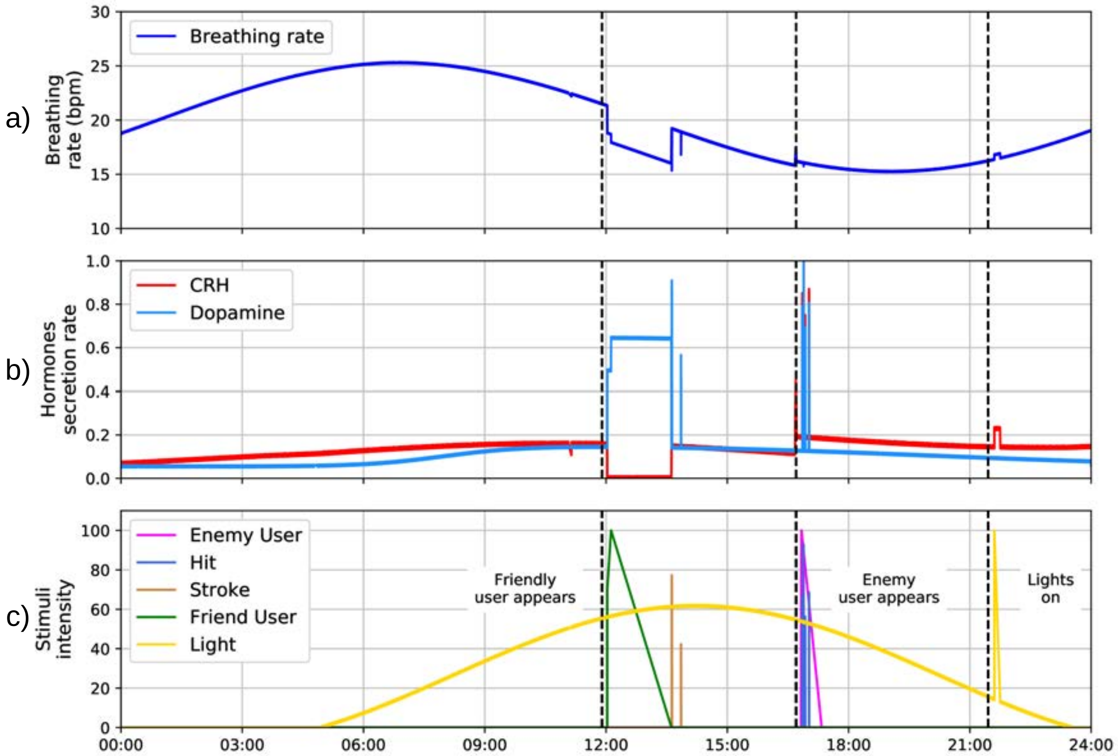


Figure 12.13: Dopamine and CRH are the two major neuromodulators of the autonomic breathing process. The breathing rate has a circadian waveform that peaks early morning and drops late evening. Many different stimuli affect the breathing rate of the robot. Positive ones evoke relaxation and a decrease in breathing per minute. Contrarily, aversive stimuli cause the opposite effect, increasing the breathing rate as a signal of nervousness.

12.2.4.5 Mini's locomotor activity

Figure 12.14 shows the locomotor activity of the social robot Mini. From left to right, the figures show different levels of locomotion. Higher locomotor activity autonomic process leads to more frequent and comprehensive movements, whereas low locomotion levels lead the robot to stay idle.

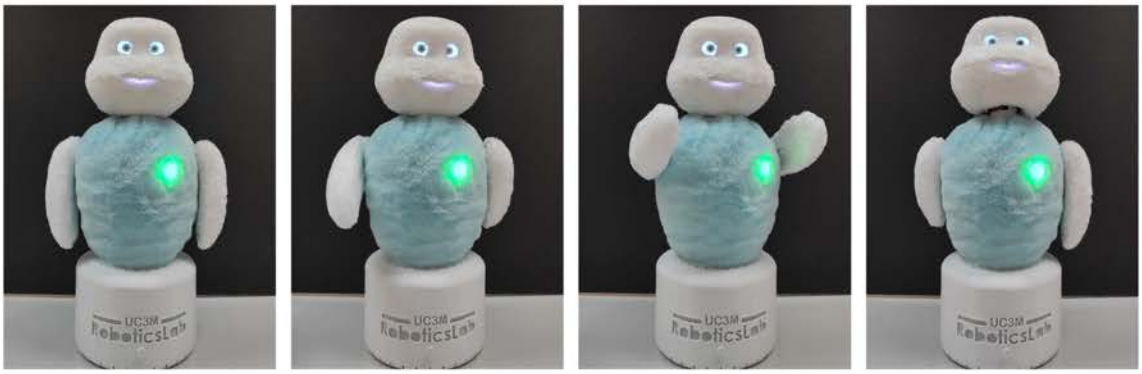


Figure 12.14: Levels of locomotor activity in the social robot Mini. From left to right, different locomotor activity values are depicted. Wide and frequent movements mean high locomotor activity levels, while an idle robot means low levels of locomotor activity.

Unlike the previous autonomic processes, locomotor activity is a process that has not been measured and standardised. For this reason, the range of the locomotor actuation autonomic process represented in the social robot Mini goes from 0 to 100 units. A value of 100 defines an aroused robot whose movements are broad in amplitude and frequent. On the opposite side, a null value of the locomotor activity defines a robot without a fluid movement. Like the blink rate, dopamine and serotonin levels promote high values of locomotor activity. However, numerous studies demonstrate that is dopamine the primary regulator of locomotion in humans.

For this reason, the locomotor activity of the robot exhibits a circadian rhythm very similar to the one exhibited by dopamine. This rhythm peaks in the afternoon (around 6pm) and drops during the dark hours. As presented in the case of the blink rate autonomic process, and depicted in Figure 12.15 (c), the perception of both positive and negative arousing stimuli rise the dopamine and serotonin levels, leading to a rapid increase in the locomotor activity of the robot.

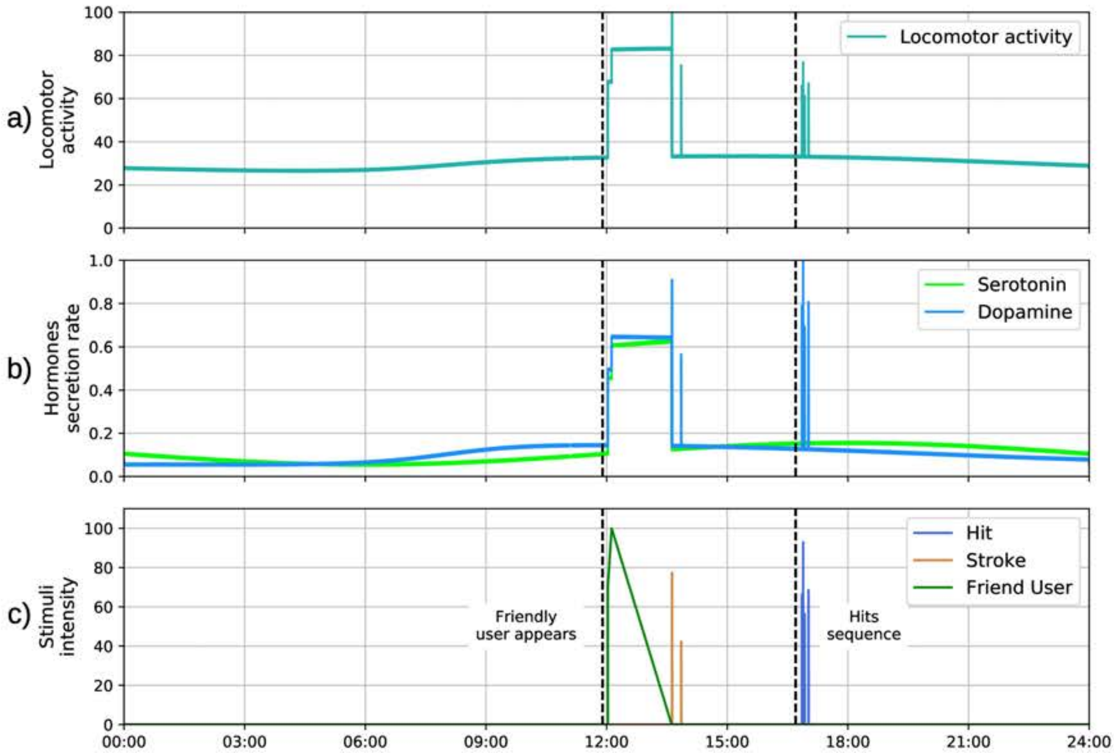


Figure 12.15: Daily evolution of the locomotor activity in the social robot Mini. Dopamine and in a lesser extent serotonin, modulate the locomotor activity of the robot. Both substances rise their levels with the perception of both positive and negative stimuli.

12.2.4.6 Simultaneous control of Mini's autonomic functions

The separate evaluation of the robot's autonomic processes yielded relevant results when analysing the cause behind each temporal evolution that these processes experienced. Nevertheless, all these processes simultaneously ran in the robot controlling its autonomic functions. Figure 12.16 shows the evolution of the heart rate, blinking frequency, pupil size, breathing rate, and locomotor activation during one natural day.

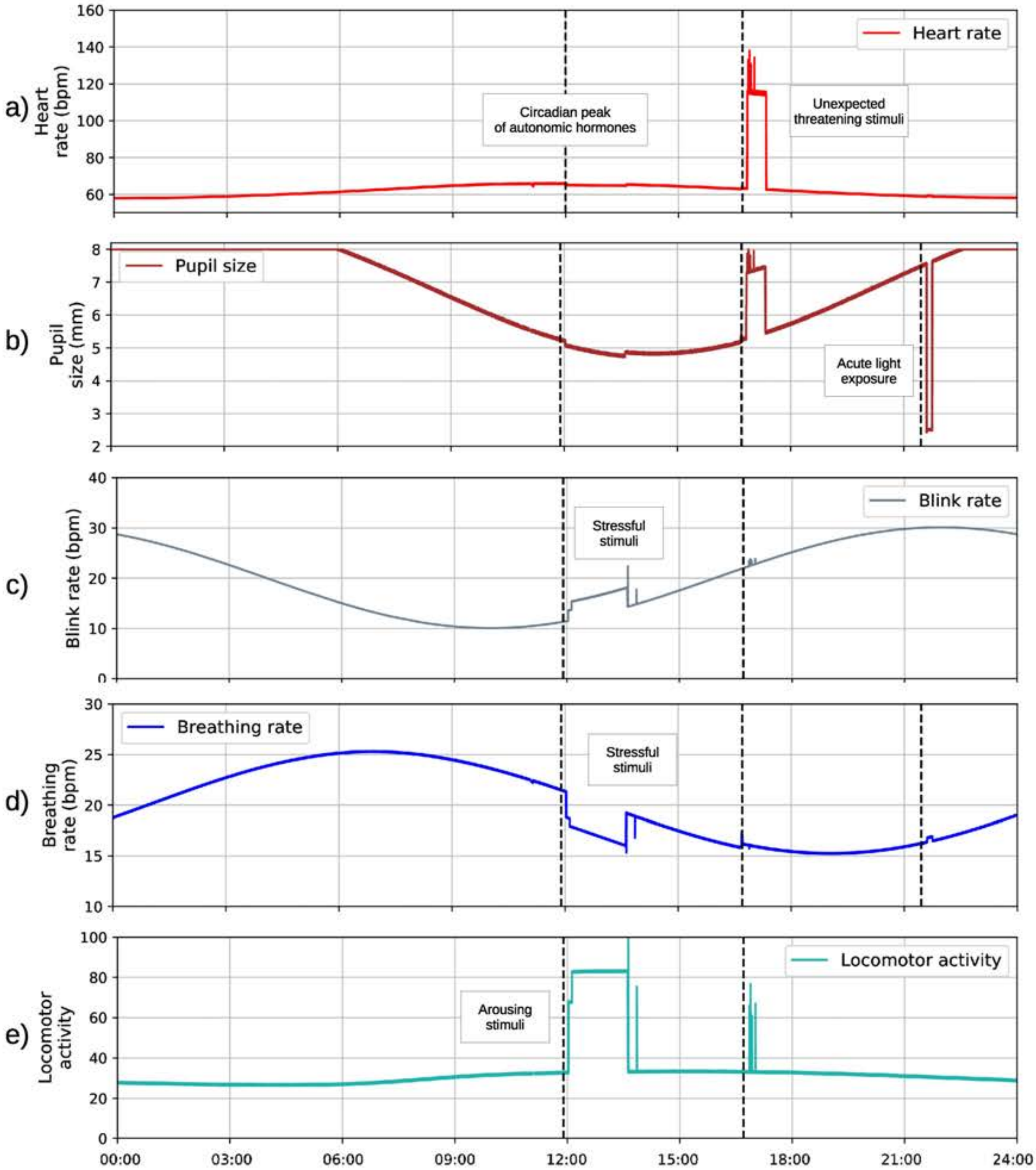


Figure 12.16: Daily evolution of Mini's autonomic functions.

- As Figure 12.16 (a) shows Mini's heart rate is a very stable signal that follows a precise circadian rhythm during the day. Nevertheless, the perception of arousing stimuli makes the heart rate rise, significantly accelerating its value and making the robot nervous and excited.

- Figure 12.16 (b) depicts the evolution of Mini's pupil size. At the beginning of the representation, Mini sleeps, so its pupil is wholly dilated (despite closed eyes). Then, with the light hours, the pupil size varies following a circadian rhythm dependent on light conditions. Besides, arousing stimuli increase pupil size and exposure to acute light contract Mini's pupils to their minimum.

- Figure 12.16 (c) shows Mini's blinking rate during a natural day. Like the previous autonomic functions, it exhibits a marked circadian pattern only affected by arousing stimuli. When the robot perceives a stimulus that makes it nervous, its blinking rate autonomously rise to show a nervous robot.

- Figure 12.16 (d) depicts Mini's breathing rate during a natural day. This autonomic process follows a circadian rhythm only altered by stressful stimuli. In these situations, the breathing rate drops, leading the robot to breath-holding. Besides, when the locomotor activity rises, the breathing rate rises slightly due to physical exercise.

- Finally, Figure 12.16 (e) shows Mini's locomotor activity. Unlike the previous processes, locomotor activity exhibits a soft circadian rhythm maintaining locomotor activity at moderate levels (around 50 units) during most of the day. However, arousing stimuli and social interaction aroused the robot, increasing its movements.

12.2.5. Discussion

Complementing the affective expressiveness of the robot, we included in the motivational model the control of physiological autonomic processes with a high impact on the robot expressiveness. In similar works, the modelling of biological processes concentrates on those involved in voluntary behaviour, leaving aside the autonomic component. Due to the lack of these studies, the modelling of autonomic responses supposes one of the main contributions of this thesis since it allows an autonomous control of physical elements of the robot as a reaction to both internal and environmental stimuli. Nonetheless, the modelling of biological processes, including autonomic ones, is restricted by the actuation capabilities of the robot. For this reason, and considering the actuators of Mini, the model regulates the heartbeat, blink rate, pupil size, respiration, and locomotor activity of the robot. Besides, the evaluation of the expressiveness of autonomic responses has not been assessed with people yet because it requires a face to face experiment that was not possible due to the world pandemic.

12.3. Conclusion

This section has presented complementary studies about affective states and autonomic physiological processes in Mini's social robot. Both works extend the previous research lines of this thesis since it goes beyond the modelling stage and the influence of these processes on decision-making. However, we believe that it is necessary not only to produce appropriate and natural behaviours during HRI but to evaluate whether the potential users of the robot correctly perceive these behaviours.

Part IV

Conclusion

The following two chapters close this thesis. First, we discuss the main results presented in the previous part, emphasising how the different functionalities of the robot work during human-robot interactions. Besides, we enumerate the limitations of our system, attempting to frame and identify future challenges in autonomous decision-making for social robots. Secondly, we conclude the dissertation providing the future works that we would like to address to continue improving the behaviour and decision-making capabilities exhibited by our social robots.

Chapter 13

Discussion and limitations

This chapter provides a general discussion about this thesis's content and main results. After that, we enumerate the principal limitations of the work presented in this dissertation, presenting possible ways to overcome them.

13.1. General discussion

Decision-Making System (DMS) are essential for machines that are intended to perform autonomous tasks. In the scenarios where autonomous machines typically work, endowing them with adaptive mechanisms are necessary to achieve their initial goals in dynamic and complex environments. In the field of robotics, and more precisely in a social context, the unpredictability of the user behaviour needs social robots to learn from past experiences anticipating new situations that may affect the correct operation of the system. This thesis has presented a DMS for social robots with adaptive and learning processes based on biologically inspired models of motivations and affect.

Nowadays, many social robots are limited to controlled scenarios with predefined and fixed tasks. Nevertheless, the necessities of society will require shortly that social robots can exhibit general and broad capabilities for accomplishing more complex tasks. Considering these facts and the intentions for the following years, social robots will, for example, be part of many different works assisting human beings in healthcare, education, or the service sectors. For this reason, and having in mind that the system introduced in this work is of general purpose, following the research line of our research group, we prioritise the design and development of software and hardware solutions for assisting caregivers during cognitive stimulation therapies with older adults presenting symptoms of diseases like Alzheimer or dementia.

Considering the workstream of our laboratory, the initial requirements of the system were precisely to develop a flexible, modular, and straightforward DMS that allows older adults to interact with social robots intuitively and, more precisely, with the social robot Mini. Both robots aim at interacting with users in one-to-one sessions. Since the robot is intended for interacting with a wide range of different users with different features, the robot has to adapt to the user in two different ways: recognising his/her limitations and facilitating the interaction process as far as possible and guiding the user's decision without interfering their real intentions and preferences.

In order to satisfy these primary goals, we proposed a modular software architecture where decisions are centralised in two hierarchical levels: the *Central controller* and several managers. The *Central controller* is in charge of receiving all the external modules of the software architecture and producing an initial decision about what action the robot should take. Then, it activates a manager that controls the execution of the action and, in some cases, makes partial decisions about the particularities of the action. The function of dividing the DMS into two levels is threefold. In the first place, it alleviates the processing of the *Central controller*, which can only focus on aggregating the information from the other modules and producing high-level decisions while the managers control the execution of the actions. Secondly, the inclusion of partial decision in the managers like, for example, the PL System avoids the *Central controlle*r to perform extra computations. Finally, it allows the system to execute multiple actions simultaneously, distributing the payload of the process. Besides, since we pretend that the robot can execute different actions simultaneously, a hierarchical system facilitates this functionality by decoupling the action control from a unique process.

One of the principal contributions of this thesis is the method for controlling the execution of the robot's actions, called *Skills*. With our method, the DMS can pause and resume ongoing actions (Skills) after evaluating the robot's situation. As stated before, the basis of this method was presented in a previous thesis of the group [19]. The actions could only be started and cancelled in the initial version, allowing their pause and resume. This functionality allows the robot's more natural and smooth behaviour since the transitions between actions do not seem as forced as just starting and stopping the activity. Besides, in this thesis, we have expanded the number of *Skills* allowing the robot to deploy a broad range of new functionalities (see Section 5.2.3 see the *Skills* of the robot).

The *Central controller* receives information from different modules of the robot. As described in Chapters 5 and 7, most of them have been developed during the realisation of this thesis to support the decision process with complete information about the robot and the environment. In this thesis, we have developed three additional modules to use the users' information and adapt the decision-making and interaction process.

- The *Context manager* is a kind of short-term memory that saves different types of data about each user interacting with the robot. The function of this module is to provide and update the user data to the rest of the modules in the software architecture, including the DMS, for producing a personalised interaction.

- The *Scheduler manager* works as the agenda of the robot, handling the events and reminders of each user stored in the *Context manager*. While reminders are short communications to the user, events are planned activities that the robot must execute during interaction with a specific user or alone.

- Finally, the robot has a *User profiling manager* that, using an intermediate robot *Skill*, allows to create from scratch an empty profile for new users that they fill during the HRI process. Besides, we included in the user profiling manager two learning algorithms related to PL focused on obtaining the preferences of each particular user to the activities that the robot can execute. All these functionalities endow the robot with a personalised behaviour towards the user.

In Chapters from 8 to 12, we assessed most of the robot's functionalities during real HRI studies not only to test the performance of each functionality but to investigate whether they improve the perception of the users about the robot in terms of usability, acceptance, intelligence perceived, and many other factors. Thus, each set of experiments reported interesting outcomes to be discussed.

- In the first place, we would like to mention the vital influence that previous research developed in the Embodied Emotion, Cognition, and Inter-Action Laboratory supposed for this work. Especially during the design and development of the DMS and the *Motivational model*. From our perspective, the speed at which artificial intelligence and machine learning are moving forward opens a disturbing but, at the same time, an extraordinary range of possibilities to carry on deepening in this fantastic world of social robotics. We believe and expect that the work presented in this thesis serves the social robots that we are working on to improve their functionalities and naturalness when interacting with people. Following these ideas, our work focused on designing a robot architecture for long-lasting adaptive and affective HRI. The outcomes we obtained demonstrated how our social robot MiRo could adapt its affective behaviour to an experimenter with different interaction limitations while promoting HRI by expressing positive emotions even when the experimenter was not interested in playing with the robot. Since social robots may meet with people who interact differently, we believe that these systems adapt to these situations and fulfil their tasks without problems.

- During the evaluation of our user profiling and PL System, we obtained interesting outcomes in each of the three stages we followed.

  - On the one hand, the evaluation of the user profiling reported that the participants preferred to create their profiles by interacting with the robot than using an online survey. Contrarily to what we initially hypothesised, the ones that interacted with a dull robot rated its usability slightly above those participants using a cheerful robot. Besides, the users creating their profile with a cheerful robot found it more engaging in entertainment than the dull robot and the online survey. In the results, it is possible to appreciate that generally, people prefer a cheerful robot. However, in the last checkpoint, the users rated above interacting with a dull robot than with a cheerful one. This fact suggests that probably, the cheerful robot may fatigue the user since the ratings of this condition decay with time, something that does not occur with the dull robot. Nevertheless, this opinion is just supposition since we do not have specific results that validate this idea.

– On the other hand, the adaptation of the robot to the user yielded meaningful results in terms of the participants' perceptions about the robot and the autonomous selections that the autonomous DMS does when working together with the PL system. First, we compared two exemplary methods for estimating the initial preferences of the user, deeming different datasets built from an online survey that retrieved the features and preferences of 471 participants. Although the results of both algorithms were excellent, we opted for integrating Ranking by PairWise Comparison (RPC) since it yielded better results. Then, we tested whether the participant considered appropriate the prediction of the system during short-term interactions. For the evaluation, we used the well-known Godspeed [348] questionnaire and six ad-hoc questions about the appropriateness of the selection and the personalisation process. Participants indicated that the activities proposed by the PL System were much more appropriate than selecting activities and random, supporting our initial beliefs. Although users found the robot presenting preferences predictions more intelligent and likeable than the one choosing activities and random, they do not perceive any difference regarding the animacy of the robot.

– Finally, these results led us to continue extending the PL System by including a second stage: adapting the initial predictions with the interaction. A simple method found in the literature suggested combining direct and indirect feedback to produce such adaptations. While the robot simply gets direct feedback by asking the user how much they liked the activity after its finalisation. Indirect feedback was more challenging to obtain since it implies the estimation of the user's likings. Taking advantage of the interaction metrics acquired from the user profiling manager for characterising the interaction, we designed a reward function that estimates how much the user likes an activity depending on how (s)he interacts with the robot while executing that action. As this was the initial testing, we opted to define a straightforward function that considers three parameters: whether the user selected the activity or ended up successfully. Results show that combining both types of feedback speeds up and stabilises the adaptation process, yielding a minor error regarding the participants' initial preferences. However, we are sure that further investigations should be addressed to improve the system. For example, the reward function can consider more metrics for estimating the indirect feedback, as well as it may be necessary to study the impact of each metric in the estimation. Besides, despite in the experiments participated 8 users during long trials, we are conscious that further testing is necessary during long-lasting real HRI scenarios. Nevertheless, we believe that producing personalised autonomous decisions is necessary for some environments, such as interacting with non-proactive users. Although some relevant works exist in the literature for endowing social robots with adaptive mechanisms, no one combines PL and RL for estimating and adapting the user preferences and personalise decision-making processes.

- Before deepening in the meaningful information that the *Motivational model* sends to the DMS to make the appropriate decision, we developed a RL method for the robot to learn how to map situations to actions. This learning scenario aimed to let the robot know its actions' effects on itself and the environment. Thus, it could autonomously behave once learned while maintaining a good internal state and surviving in a dynamic environment. The two experiments we designed (Q-learning and Dyna-Q+) allowed the robot to attain its goals and exhibit an autonomous and coherent behaviour by recognising the effects of its actions. However, Q-learning needed very long training periods, so we used Dyna-Q+ (a model-based approach) to overcome the Q-learning issues. Our results 'demonstrated that Dyna-Q+ is a powerful technique that improves Q-learning in learning speed and stability but requires a precise model of the environment and more computational load.

- As presented in Chapter 6, probably the most important contribution together with the DMS system is the *Motivational model*. In the last decades, many researchers have been working in this are attempting to endow a robot with biologically inspired capabilities to produce reasonable and appropriate decisions. Despite this thesis taking powerful inspiration from these studies, we have extended them by including important contributions in our *Motivational model* in neuroscience. The *Motivational model* introduced in this manuscript considers neuroendocrine responses as the link between external stimuli and physiological and psychological processes that are the basis of motivated behaviour. For this reason, the robot has several artificial hormones and neurotransmitters whose levels determine the evolution of multiple voluntary and autonomic biological processes in the robot. Besides, the model avoids modelling the evolution of these processes using linear functions since many studies provide robust background about the reactivity and non-linearity of living organisms. The model also presents outstanding notes about the generation and expression of the affective state, which has been represented by combining emotional and mood responses that can overlap and compete among them. Another important contribution shown in the model is the differentiation between voluntary and involuntary behaviour and autonomic responses. On the one hand, we model voluntary behaviour as proactive motivations whose intensity majorly depends upon the deficits of physiological processes.

  On the other hand, involuntary behaviour is a reactive motivational state that suddenly responds to unexpected stimuli. In the model, a proactive motivation and a reactive one can blend during seconds, improving the naturalness of the system when reacting to unforeseen situations. From our point of view, our *Motivational model* expands previous biologically inspired models giving more capabilities to social robots. Nevertheless, we also think that we are still far from obtaining 24-hours robots in uncontrolled environments because generating complex adaptive behaviour requires a very precise adjustment of the emulated biological processes. Consequently, and according to our experience, the main problem of the *Motivational model* is the characterisation of each process. Currently, a method does not exist to map the evolution of human internal processes into mathematical expressions that entirely define these rhythms during long periods. For this reason, the attributes defining each process have to be empirically set and adjusted using trial and error until the system attains a final optimal solution. Thus, modelling a large number of processes would become very tedious. Considering these ideas, the number of processes in our model is still far from the real amount of processes inside a natural organism, especially if considering psychological aspects of human behaviour. Our idea is to continue working with our *Motivational model* as we detail in the Future work Section 14.2.

- Finally, we deepened into the expressive aspect of social robots in affect and the autonomic control of physiological processes. Thus, we wanted to endow social robots with reactive involuntary behaviour, a field that we perceived is not as explored as proactive behaviour. On the one hand, since the *Motivational model* generates the robot's affective state, we sought to design affective expressions to let the users know how Mini feels. Since these expressions were manually adjusted, we evaluated what people thought about them in a preliminary study. Once we obtained affective expressions that people could correctly perceive, we designed a second scenario where the affective modulation was carried out in different ways by combining punctual emotions, the decay of such emotions with time, and mood. Our initial hypothesis posited that people would better perceive a robot showing punctual emotions, decay with time, and mood. However, our hypothesis was incorrect since people indicated that the easiest way to recognise emotions was when the robot just expressed punctual emotional reactions.

  On the other hand, we believe that for social robots to fully exhibit autonomous and natural behaviours, endowing them with autonomic control is necessary to provide good expressiveness and let the user know what the robot is experiencing. For this reason, and taking strong inspiration from various studies in neuroscience, we empirically characterised the evolution of the heart rate, the blink rate, the breathing rate, the pupil size, and the locomotor activation of our social robots according to neuroendocrine responses. Thus, embodied representation of these processes evolves following circadian rhythms and the impact of certain neuromodulators like epinephrine, norepinephrine, dopamine, or serotonin.

To conclude with this general discussion, we would like to stress the importance of DMS for autonomous robots. This research line is still in a current stage that requires more work and investigation to attain fully autonomous systems in complex and dynamic scenarios. In the following section, we present our system's limitations from a computational and conceptual point of view, giving some ideas about dealing with such disadvantages.

13.2. Limitations

Despite the initial objectives of this thesis, there are still many issues to address to deploy our social robots and DMS in long-term interactions scenarios. The following paragraphs define the main limitations of this thesis, giving possible solutions that may help the reader to understand our plans to overcome them.

- **Fine tuning of parameters:** One of the main contributions of this work is to present a *Motivational model* that allows the DMS to make decisions using biologically inspired foundations. However, like Chapter 6 shows, the model requires the tuning of a wide range of parameters that define how the internal system of the robot work. Most of these parameters have been empirically set by trial and error to obtain a specific temporal evolution of the artificial biological processes of the robot. In fact, due to the large set of parameters that the model has, in some of the relationships between them (like in Table 6.5), we had to generalise the values of the parameters clustering them in three levels (low, moderate, and strong) to simplify the modelling process. At this moment, solving this issue is not an easy task. Like occurs in machine learning techniques with the tuning of the hyperparameters, a possible solution can be to use generalisation methods like defining different levels of interactions. However, due to the particularity of our system, some processes will continue requiring custom tuning of their parameters.

- **Formal representation of biological processes:** In line with the previous problem, and despite the many exciting advances in neuroscience and modelling research, there is still a lack of mathematical equations regarding the evolution of our internal processes. The internal process of our organism is a complex system with a considerable degree of dependability. This fact, summed up to the differences between individuals, makes modelling this process an arduous task. To address this issue, combining different research areas like medicine or psychology with robotics is necessary to improve the mathematical equations behind biology.

- **DMS flexibility:** The DMS is the cornerstone of this thesis. It allows our social robots to make autonomous decisions merging information from many diverse modules like the scheduler, the *Motivational model*, or the user profiling. However, the messages received by each input of the DMS are not homogeneous. Because of this, the DMS needs to include specific functions for reading and understanding the information that each module sends. This problem leads to a decrease in the flexibility and computation speed of the program, causing the system to work slowly in situations where the decision clusters much external information. For this reason, the idea for the future is to standardise the data that the DMS receives for simplifying the communication in the software architecture.

- **Dimensionality problem in RL:** The results shown in Chapters 8 and 9 demonstrate how RL is a powerful technique for learning while interacting with the environment that surrounds the robot. Nevertheless, RL contains many limitations. The most outstanding limitation of RL tabular methods is the negative effect that large state-action spaces produce in the learning speed and memory storing. A possible solution to this issue is to use deep learning and function approximation at the cost of increasing the computation of the system and the computational expenditure, especially in real-time applications like social robotics. The learning speed is also massively dependent on the duration of the action. Thus, since RL methods require repeated interactions with the environment to converge to an optimal solution, the combination of a large action space with long actions can be intractable. Model-based methods like Dyna-Q are a good alternative for reducing the effect of large state-action spaces, as they allow planning in the background.

- **Shaping reward functions:** Nonetheless, the model requires to be very well-defined to represent the real world accurately. Otherwise, it can become a double-edged sword. The solution we propose in the thesis is a model of the environment that allows to speed up the learning process but includes an adaptive mechanism that allows the model to self-improve by acknowledging the changes in the environment. Besides, the design of the rewards functions marks the success of the learning system since they exactly represent what the agent is learning. Despite reward functions are not fully considered as a hyperparameter, it is dependent on the application where the RL system works and is decided by the designer. In this thesis, we use two different rewards functions in different problems. Despite both yielded very positive outcomes, it is impossible to know the best possible alternatives. For this reason, the reward function themselves suppose a significant drawback in RL problems.

- **Problems in preference prediction:** The PL system presented in this thesis allows the robot to suggest to each user their favourite entertaining activities when interacting with the robot. Although the evaluation with real users yielded very positive results, the learning system has some limitations that should be considered. In the first place, the system's performance is highly dependent on the number of labels to rank. A possible solution for these issues is to decompose the learning scenario into minor problems where the number of labels to rank decreases. Despite these solutions will probably increase the performance of the learning system, it may increase the computational time required to make predictions since the number of datasets increases. As mentioned previously, the performance of this kind of system is susceptible to the dataset's structure and its building. To overcome this problem, it is necessary to increase the number of dataset instances to improve the system's performance while decreasing the sensitivity to cross-validation. Finally, a well-known drawback of machine learning techniques is also present in our system. Hyperparameter optimisation is a fundamental problem in the performance of the algorithms. In this work, we use a grid search method [347] to find the most appropriate hyperparameters of the system, although this method increases the training time of the mode.

- **Covid-19 pandemic:** Due to the Covid-19 pandemic, the evaluation of the functionalities developed in this thesis during face-to-face human-robot interactions were evaluated in many cases using online surveys. Despite it being a practical method in some situations, we believe that face-to-face interactions are a must for users to feel the ins and outs of the robot. For this reason, the evaluation of the experiments presented in this dissertation combines online evaluations with real interactions. In fact, in most of the cases presented in the previous chapter, we attempted to carry out the tests face-to-face but the assessing using online questionnaires.

- **Definition and modelling of affect:** The terminology behind affect theories is diffuse and sparse in some cases. Although most theories agree in defining the concepts of emotion, mood, and personality, serving us as a great source of inspiration, some theories do not agree on where to set the line that differentiates such concepts. Besides, the influence of affective states in behaviour is a research line that, from our point of view, requires further investigation due to the intrinsic complexity that it comes with. This problem arose in this thesis during the realisation of the experiments presented in Chapter 12, focused on evaluating the affective expressiveness of the social robot Mini. The results show that the participants could correctly perceive the individual expression of the robot's emotional and mood states. However, they could not perceive any differences in the robot expressing complex affective regulation compared to a robot showing limited affective expressiveness. These results suggest that further experiments are still necessary to endow a robot with affective capabilities.

- **Lack of a general evaluation:** The results presented in this thesis comprise the individual evaluations of the different functionalities of the system. However, the reader may miss a detailed user study about the evaluation of the whole system integrating the diverse functionalities assessed individually. Due to the Covid-19 pandemic and the current state of our robots, we believe that the evaluation of the whole system is not possible yet, even though we would like to carry out this experiment as soon as possible. Our idea is to deploy some of our robots in nursing centres to interact with older adults during long periods. In this way, we will test the robot's performance to measure its usability and participants' acceptance.

CHAPTER 14

Conclusions and future work

In this chapter, we provide the main conclusions of this work, finishing with the future work that we will tackle in order to continue improving our DMS.

14.1. Conclusions

In this dissertation, we have described a DMS devised to allow social robots to behave autonomously in a dynamic environment. The system receives inputs from various modules to produce a biologically inspired behaviour that can adapt to the user interacting with the robot. The robot can learn from the interaction with the environment the best actions it can execute in every particular situation, leading it to maintain its physiological and psychological state in good condition. Besides, the robot can interact with people adapting to each specific user by creating social bonds with him/her and learning their preferences to the entertaining activities that it can execute. Thus, the PL System predicts and adapts the user preferences depending on his/her selections and actions, improving the communication between both agents.

Like the results presented in Part III show, the evaluation of the system was carried out in different stages instead of assessing the entire operation of the system.

- The first experiments we presented in this manuscript were oriented to endow social robots with adaptive behaviour to improve HRI. At the same time, the robot fostered the expression of positive affective emotions to engage users to overcome their interaction limitations. This work, developed at the University of Hertfordshire, was an important influence for the rest of this thesis since it provided us with new ideas for our adaptive and biologically inspired DMS.
- Secondly, we showed how the robot could adapt its behaviour to the user in two different areas: adapting the interaction by including the user's personal information and estimating and learning each particular user's preferences towards suggesting their preferred entertainment activities. Results demonstrate that a robot able to adapt its behaviour to the user is more accepted and perceived as intelligent than a robot that does not present this capability. Besides, the user likes the

robot in general terms. However, as expected, any differences are obtained in anthropomorphism, animacy, or security perceived.

- The third group of experiments pretended to assess the operation of the *Motivational model* in two different scenarios: testing that the artificial biological processes of the robot follow a reasonable daily evolution and whether the robot can learn which behaviour suits better in which situation where it is involved. The evolution of the biological processes proves that the robot can deploy a natural behaviour during human-robot interaction while satisfying their internal needs. Besides, the robot can affectively react, expressing its emotions and mood. Results show how people successfully perceived the robot's affect. However, as mentioned in the limitations of this architecture, there is still much work to be done to reach a modelling method that allows generalising each biological process included in the system. It is interesting how circadian variations of the neurotransmitters dopamine, norepinephrine, and serotonin modulate affective responses, causing that, for example, the mood lowers in the early morning, improving along the day.

- Additionally, this thesis addresses mapping specific biological processes into body expressions and implementing artificial biological processes in embodied systems. Specifically, we concentrate on expressing emotional responses, moods, and autonomic processes to improve how the users perceive the robot. Emotions are reactive affective processes to situations that the robot unexpectedly experiences. The effect of emotional responses rapidly decreases in time, being strongly affected by the perception of certain stimuli. The robot's mood is a long-term affective state derived from the accumulation of past emotional experiences and the physiological state of the robot. Unlike emotions, the mood is not a reactive process tied to specific stimuli but long-lasting affective states whose effect disappears without arousing stimuli. As the results show, users consider more positively affective robots than neutral ones since they perceive them as more natural and expressive.

- Although affective states improve the expressiveness and naturalness of robots, little attention has been paid to the expression of autonomic processes. Considering the lack of research in modelling and regulating autonomic processes in social robots from a biologically inspired approach, we studied how to regulate the autonomic processes of the social robot Mini from a neuroendocrine perspective. Results show how specific hormones and neurotransmitters modulate the heart rate, blink rate, pupil size, breathing rate, and locomotion activation. These autonomic processes vary depending on internal circadian rhythms, the interaction between chemicals, and external stimuli' strong influence. While the circadian rhythms and the interaction between chemicals suppose the base of the evolution that each autonomic process follows, the influence of external stimuli supposes the reactive component to maintaining good internal conditions.

Consequently, the purpose of regulating the robot's autonomic processes is twofold. First, it translates arousing stimuli into autonomous expressions that people can easily perceive and interpret since human responses inspire them. Second, the biologically inspired modelling provides a natural and well-grounded definition of the processes, making their evolution reasonable and meaningful.

This thesis shows that social robotics, artificial intelligence, and biologically inspired DMSs are growing by leaps and bounds. The number of opportunities that these topics open in areas like healthcare, education, or assistance makes us enormously motivated to continue improving this system. Hence, the following section enumerates future steps to improve the system presented in this thesis prevailing social robots' autonomy, adaptation, and intelligence.

14.2. Future work

Despite the notable improvements that we have obtained in endowing our social robots with autonomous decision-making for adapting the interaction with the user, there are still many areas to research and many issues to overcome. The following paragraphs enumerate the next steps to improve the system presented in this dissertation and overcome the limitations presented in the previous chapter.

- **Deployment in real scenarios:** The first and most important future work is to evaluate the operation of the architecture in long-term studies with different kinds of populations (e.g. children, older adults), integrating all the functionalities described in this manuscript. The evaluations were conducted in separate experiments to test each robot's functionality. However, testing the full system together will provide us with an accurate view of the real needs of people when facing and interacting with a social robot. The results obtained will drive the incorporation of new functionalities according to the fundamental requirements of robot users. We plan to address this task in the following months building new robots and deploying them in nursing homes.
- **Behaviour simultaneity:** Human behaviour is diverse and unpredictable. Moreover, humans can execute multiple actions simultaneously, paying attention to multiple changes around us and reacting to them. Considering this idea, the current state of the DMS does not allow to execute multiple actions at the same time, as the current active activity has to be paused or stopped to execute a new one. Thus, the system can be notably improved by including the possibility to combine different behaviours at the same time in specific circumstances. This functionality will allow the robot to perform a more natural and varied behaviour, something that we believe will produce very positive outcomes in terms of usability and acceptability.

- **Emulate new biologically inspired processes:** The inclusion of a *Motivational model* provides a wide range of novel and exciting areas to research. Considering this, we would like to continue exploring and modelling those human processes involved in decision-making like curiosity or obedience. We believe that curiosity will lead the robot to explore its environment producing varied behaviour. Besides, modelling obedience will provide the robot with mechanisms that allow it to balance between the user petitions and its own needs.

- **Expand the robot's affective states:** Currently, the model generates 4 emotions and 9 moods. The idea is to continue expanding the affective states that the robot can express by researching the effects of emotions like fear on the robot's behaviour.

- **Explore new learning methods:** Machine learning is progressing at a frenetic pace. The appearance of new techniques and algorithms and their application to social robotics will provide us with fantastic tools to continue improving self-learning and adaptation capabilities. Despite focusing on adapting to the user right now, our next steps will include learning mechanisms for learning how to interpret new stimuli or dealing with multi-user domains.

- **Multi-user interaction:** Considering the previous future line, our system allows the robot to interact with one user simultaneously. The system can dynamically adapt its behaviour to the user it interacts with but can not concentrate on more than one user simultaneously. For this reason, we would like to study how to deal with situations in which the robot perceives more than one user simultaneously, adapting the interaction to all of them simultaneously.

- **Improve robot personalisation:** The PL System presented in this thesis comprises two steps: prediction and online adaptation. Despite the current state of the system works properly, allowing the robot to suggest each participant with their favourite entertaining activities, the idea is to continue improving moving one step further in optimising the initial preference estimation provided by the LR model. As stated in Chapter 9, right now, the preference estimation comes from a predefined dataset built using an online survey. The adapted preferences of the users that interacted with the robot will serve to feedback the database by adding new instances that can improve new user preference predictions. Since this new function requires retrieving many different user attributes, we pretend to fuse the PL System with active user profiling to reinforce the preferences predicted by the model.

Bibliography

[1] Dean Mobbs, Pete C Trimmer, Daniel T Blumstein and Peter Dayan. 'Foraging for foundations in decision neuroscience: insights from ethology'. In: *Nature Reviews Neuroscience* 19.7 (2018), pp. 419–427.

[2] Thea Ionescu and Dermina Vasc. 'Embodied cognition: challenges for psychology and education'. In: *Procedia-Social and Behavioral Sciences* 128 (2014), pp. 275–280.

[3] Sébastien Tremblay, KM Sharika and Michael L Platt. 'Social decision-making and the brain: A comparative perspective'. In: *Trends in cognitive sciences* 21.4 (2017), pp. 265–276.

[4] Rebecca A. Ferrer and Wendy Berry Mendes. 'Emotion, health decision making, and health behaviour'. In: *Psychology & Health* 33.1 (2018), pp. 1–16.

[5] Valeria Villani, Fabio Pini, Francesco Leali and Cristian Secchi. 'Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications'. In: *Mechatronics* 55 (2018), pp. 248–266.

[6] Brian Paden, Michal Čáp, Sze Zheng Yong, Dmitry Yershov and Emilio Frazzoli. 'A survey of motion planning and control techniques for self-driving urban vehicles'. In: *IEEE Transactions on intelligent vehicles* 1.1 (2016), pp. 33–55.

[7] Xueman Fan, J Wu and L Tian. 'A review of artificial intelligence for games'. In: *Artificial Intelligence in China* (2020), pp. 298–303.

[8] Gabriela Prelipcean, Mircea Boscoianu and Florin Moisescu. 'New ideas on the artificial intelligence support in military applications'. In: *Proceedings of the 9th WSEAS international conference on Artificial intelligence, knowledge engineering and data bases*. World Scientific, Engineering Academy and Society (WSEAS). 2010, pp. 34–39.

[9] Pavel Hamet and Johanne Tremblay. 'Artificial intelligence in medicine'. In: *Metabolism* 69 (2017), S36–S40.

[10] Weiyu Wang and Keng Siau. 'Artificial intelligence, machine learning, automation, robotics, future of work and future of humanity: A review and research agenda'. In: *Journal of Database Management (JDM)* 30.1 (2019), pp. 61–79.

[11] Marco Nørskov. *Social robots: boundaries, potential, challenges*. Taylor & Francis, 2017.

[12] Ronald Craig Arkin, Patrick Ulam and Alan R Wagner. 'Moral decision making in autonomous systems: Enforcement, moral emotions, dignity, trust, and deception'. In: *Proceedings of the IEEE* 100.3 (2011), pp. 571–589.

[13] Yingxu Wang and Guenther Ruhe. 'The cognitive process of decision making'. In: *International Journal of Cognitive Informatics and Natural Intelligence (IJCINI)* 1.2 (2007), pp. 73–85.

[14] Miguel A Salichs et al. 'Maggie: A robotic platform for human-robot social interaction'. In: *2006 IEEE conference on robotics, automation and mechatronics*. IEEE. 2006, pp. 1–7.

[15] María Ángeles Malfaz Vázquez. 'Sistema de toma de decisiones basado en emociones y autoaprendizaje para agentes sociales autónomos'. PhD thesis. University Carlos III of Madrid, 2007.

[16] Chris Watkins and Peter Dayan. 'Q-learning'. In: *Machine Learning* 8.3 (May 1992), pp. 279–292.

[17] Álvaro Castro González. 'Bio-inspired decision making system for an autonomous social robot: the role of fear'. PhD thesis. University Carlos III of Madrid, 2012.

[18] Fernando Alonso Martín. 'Sistema de interacción humano-robot basado en diálogos multimodales y adaptables'. PhD thesis. University Carlos III of Madrid, 2014.

[19] Raúl Pérula-Martínez. 'Autonomous decision-making for socially interactive robots'. PhD thesis. University Carlos III of Madrid, 2017.

[20] Filippo Cavallo, Francesco Semeraro, Laura Fiorini, Gergely Magyar, Peter Sinčák and Paolo Dario. 'Emotion modelling for social robotics applications: a review'. In: *Journal of Bionic Engineering* 15.2 (2018), pp. 185–203.

[21] Arielle AJ Scoglio, Erin D Reilly, Jay A Gorman and Charles E Drebing. 'Use of social robots in mental health and well-being research: systematic review'. In: *Journal of medical Internet research* 21.7 (2019), e13322.

[22] Jonathan Casas, Nathalia Cespedes, Marcela Múnera and Carlos A Cifuentes. 'Human-robot interaction for rehabilitation scenarios'. In: *Control Systems Design of Bio-Robotics and Bio-mechatronics with Advanced Applications*. Elsevier, 2020, pp. 1–31.

[23] Mohammed A Saleh, Fazah Akhtar Hanapiah and Habibah Hashim. 'Robot applications for autism: a comprehensive review'. In: *Disability and Rehabilitation: Assistive Technology* 16.6 (2021), pp. 580–602.

[24] VA Golovko. 'Deep learning: an overview and main paradigms'. In: *Optical memory and neural networks* 26.1 (2017), pp. 1–17.

[25] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

[26] Philippe Aghion, Benjamin F Jones and Charles I Jones. *Artificial Intelligence and Economic Growth*. Working Paper 23928. National Bureau of Economic Research, Oct. 2017.

[27] Arshia Khan and Yumna Anwar. 'Robots in healthcare: A survey'. In: *Science and Information Conference*. Springer. 2019, pp. 280–292.

[28] Tony Belpaeme, James Kennedy, Aditi Ramachandran, Brian Scassellati and Fumihide Tanaka. 'Social robots for education: A review'. In: *Science robotics* 3.21 (2018).

[29] Eloise Matheson, Riccardo Minto, Emanuele GG Zampieri, Maurizio Faccio and Giulio Rosati. 'Human–robot collaboration in manufacturing applications: a review'. In: *Robotics* 8.4 (2019), p. 100.

[30] Andrea Bonarini. 'Communication in Human-Robot Interaction'. In: *Current Robotics Reports* (2020), pp. 1–7.

[31] Lihui Pu, Wendy Moyle, Cindy Jones and Michael Todorovic. 'The effectiveness of social robots for older adults: a systematic review and meta-analysis of randomized controlled studies'. In: *The Gerontologist* 59.1 (2019), e37–e51.

[32] Carlos A Cifuentes, Maria J Pinto, Nathalia Céspedes and Marcela Múnera. 'Social robots in therapy and care'. In: *Current Robotics Reports* (2020), pp. 1–16.

[33] Yaoxin Zhang et al. 'Could social robots facilitate children with autism spectrum disorders in learning distrust and deception?' In: *Computers in Human Behavior* 98 (2019), pp. 140–149.

[34] Martina Čaić, Dominik Mahr and Gaby Oderkerken-Schröder. 'Value of social robots in services: social cognition perspective'. In: *Journal of Services Marketing* (2019).

[35] Paul Formosa. 'Robot Autonomy vs. Human Autonomy: Social Robots, Artificial Intelligence (AI), and the Nature of Autonomy'. In: *Minds and Machines* (2021), pp. 1–22.

[36] Richelle ACM Olde Keizer et al. 'Using socially assistive robots for monitoring and preventing frailty among older adults: a study on usability and user experience challenges'. In: *Health and Technology* 9.4 (2019), pp. 595–605.

[37] Nick Gravish and George V Lauder. 'Robotics-inspired biology'. In: *Journal of Experimental Biology* 221.7 (2018), jeb138438.

[38] Omar Mubin, Muneeb Imtiaz Ahmad, Simranjit Kaur, Wen Shi and Aila Khan. 'Social robots in public spaces: A meta-review'. In: *International Conference on Social Robotics*. Springer. 2018, pp. 213–220.

[39] Sebastian Thrun. 'Probabilistic algorithms in robotics'. In: *Ai Magazine* 21.4 (2000), pp. 93–93.

[40] Tobias Kaupp, Alexei Makarenko and Hugh Durrant-Whyte. 'Human–robot communication for collaborative decision making—A probabilistic approach'. In: *Robotics and Autonomous Systems* 58.5 (2010), pp. 444–456.

[41] Michał Czubenko, Zdzisław Kowalczuk and Andrew Ordys. 'Autonomous driver based on an intelligent system of decision-making'. In: *Cognitive computation* 7.5 (2015), pp. 569–581.

[42] Jacaueline Hemminahaus and Stefan Kopp. 'Towards adaptive social behavior generation for assistive robots using reinforcement learning'. In: *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI*. IEEE. 2017, pp. 332–340.

[43] Amir Ramezani Dooraki and Deok Jin Lee. 'Memory-based reinforcement learning algorithm for autonomous exploration in unknown environment'. In: *International Journal of Advanced Robotic Systems* 15.3 (2018), p. 1729881418775849.

[44] Hae Won Park, Ishaan Grover, Samuel Spaulding, Louis Gomez and Cynthia Breazeal. 'A model-free affective reinforcement learning approach to personalization of an autonomous social robot companion for early literacy education'. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 33. 01. 2019, pp. 687–694.

[45] John Lones, Matthew Lewis and Lola Cañamero. 'A hormone-driven epigenetic mechanism for adaptation in autonomous robots'. In: *IEEE Transactions on Cognitive and Developmental Systems* 10.2 (2017), pp. 445–454.

[46] Kingson Man and Antonio Damasio. 'Homeostasis and soft robotics in the design of feeling machines'. In: *Nature Machine Intelligence* 1.10 (2019), pp. 446–452.

[47] Vasiliki Vouloutsi and PFMJ Verschure. 'Emotions and self-regulation'. In: *Living Machines, a Handbook of Research in Biomimetic and Biohybrid Systems* (2018), pp. 327–337.

[48] JM Fuster. 'Prefrontal cortex in decision-making: The perception–action cycle'. In: *Decision neuroscience*. Elsevier, 2017, pp. 95–105.

[49] Chris D Frith and Uta Frith. 'Social cognition in humans'. In: *Current biology* 17.16 (2007), R724–R732.

[50] Muneeb Imtiaz Ahmad, Omar Mubin and Joanne Orlando. 'Adaptive social robot for sustaining social engagement during long-term children–robot interaction'. In: *International Journal of Human–Computer Interaction* 33.12 (2017), pp. 943–962.

[51] Silvia Rossi, Francois Ferland and Adriana Tapus. 'User profiling and behavioral adaptation for HRI: A survey'. In: *Pattern Recognition Letters* 99 (2017), pp. 3–12.

[52] Tsung-Ren Huang et al. ‘Asynchronously Embedding Psychological Test Questions into Human–Robot Conversations for User Profiling’. In: *International Journal of Social Robotics* 13.6 (2021), pp. 1359–1368.

[53] Ning Wang, Alessandro Di Nuovo, Angelo Cangelosi and Ray Jones. ‘Temporal patterns in multi-modal social interaction between elderly users and service robot’. In: *Interaction Studies* 20.1 (2019), pp. 4–24.

[54] Luis Cobo Hurtado, Pablo Francisco Viñas, Eduardo Zalama, Jaime Gómez-García-Bermejo, José María Delgado and Beatriz Vielba García. ‘Development and Usability Validation of a Social Robot Platform for Physical and Cognitive Stimulation in Elder Care Facilities’. In: *Healthcare*. Vol. 9. 8. Multidisciplinary Digital Publishing Institute. 2021, p. 1067.

[55] Cynthia Breazeal, Kerstin Dautenhahn and Takayuki Kanda. ‘Social robotics’. In: *Springer handbook of robotics* (2016), pp. 1935–1972.

[56] Sally Whelan, Kathy Murphy, Eva Barrett, Cheryl Krusche, Adam Santorelli and Dympna Casey. ‘Factors affecting the acceptability of social robots by older adults including people with dementia or cognitive impairment: a literature review’. In: *International Journal of Social Robotics* 10.5 (2018), pp. 643–668.

[57] Minoo Alemi, Alireza Taheri, Azadeh Shariati and Ali Meghdari. ‘Social Robotics, Education, and Religion in the Islamic World: An Iranian Perspective’. In: *Science and Engineering Ethics* 26.5 (2020), pp. 2709–2734.

[58] Carlo Cantile and Sameh Youssef. ‘Nervous system’. In: *Jubb, Kennedy & Palmer’s Pathology of Domestic Animals: Volume 1* (2016), p. 250.

[59] Per Brodal. *The central nervous system: structure and function*. Oxford University Press, 2004.

[60] Laurie Kelly McCorry. ‘Physiology of the autonomic nervous system’. In: *American journal of pharmaceutical education* 71.4 (2007).

[61] Micky A Akinrodoye and Forshing Lui. ‘Neuroanatomy, Somatic Nervous System’. In: *StatPearls Publishing LLC* (2020).

[62] K Ondicova and B Mravec. ‘Multilevel interactions between the sympathetic and parasympathetic nervous systems: a minireview’. In: *Endocr Regul* 44.2 (2010), pp. 69–75.

[63] Rae Silver and Lance J Kriegsfeld. ‘Hormones and behaviour’. In: *e LS* (2001).

[64] Walter Bradford Cannon. ‘Homeostasis’. In: *The wisdom of the body* (1932).

[65] Steven J Cooper. ‘From Claude Bernard to Walter Cannon. Emergence of the concept of homeostasis’. In: *Appetite* 51.3 (2008), pp. 419–427.

[66] Vladyslav V Vyazovskiy, Mark E Walton, Stuart N Peirson and David M Bannerman. ‘Sleep homeostasis, habits and habituation’. In: *Current opinion in neurobiology* 44 (2017), pp. 202–211.

[67] Elizabeth A Lawson, Pawel K Olszewski, Aron Weller and James E Blevins. 'The role of oxytocin in regulation of appetitive behaviour, body weight and glucose homeostasis'. In: *Journal of neuroendocrinology* 32.4 (2020), pp. 1–19.

[68] Theodor H Benzinger. 'Heat regulation: homeostasis of central temperature in man'. In: *Physiological reviews* 49.4 (1969), pp. 671–759.

[69] Miriam A Bredella. 'Sex differences in body composition'. In: *Sex and gender factors affecting metabolic homeostasis, diabetes and obesity* (2017), pp. 9–27.

[70] D Weinert and DG Gubin. 'The circadian body temperature rhythm-Origin and implications for health and wellbeing'. In: *Tyumen Medical Journal* 20.2 (2018).

[71] Jay Schulkin and Peter Sterling. 'Allostasis: a brain-centered, predictive mode of physiological regulation'. In: *Trends in neurosciences* 42.10 (2019), pp. 740–752.

[72] Jay Schulkin et al. *Allostasis, homeostasis, and the costs of physiological adaptation*. Cambridge University Press, 2004.

[73] Jay Schulkin. 'Social allostasis: anticipatory regulation of the internal milieu'. In: *Frontiers in evolutionary neuroscience* 2 (2011), p. 111.

[74] Darby E Saxbe, Lane Beckes, Sarah A Stoycos and James A Coan. 'Social allostasis and social allostatic load: a new model for research in social dynamics, stress, and health'. In: *Perspectives on Psychological Science* 15.2 (2020), pp. 469–482.

[75] MM Reshetnikov. 'Problem of relation between brain and mind in physiology, medicine and psychology'. In: *Journal of Psychiatry and Psychiatric Disorders* 1.6 (2017), pp. 313–316.

[76] Philip E Cryer. 'Physiology and pathophysiology of the human sympathoadrenal neuroendocrine system'. In: *New England Journal of Medicine* 303.8 (1980), pp. 436–444.

[77] John E Hall and Michael E Hall. *Guyton and Hall textbook of medical physiology e-Book*. Elsevier Health Sciences, 2020.

[78] C Ronald Kahn. 'Membrane receptors for hormones and neurotransmitters.' In: *Journal of Cell Biology* 70.2 (1976), pp. 261–286.

[79] J Knight. 'Endocrine system I: overview of the endocrine system and hormones'. In: *Nursing Times* (2021), pp. 38–42.

[80] Bo Si and Edward Song. 'Recent advances in the detection of neurotransmitters'. In: *Chemosensors* 6.1 (2018), p. 1.

[81] Tyler J Stevenson, Haley E Hanson and Lynn B Martin. 'Theory, hormones and life history stages: an introduction to the symposium epigenetic variation in endocrine systems'. In: *Integrative and Comparative Biology* 60.6 (2020), pp. 1454–1457.

[82] Irina V Zhdanova and Valter Tucci. ‘Melatonin, circadian rhythms, and sleep’. In: *Current treatment options in neurology* 5.3 (2003), pp. 225–229.

[83] C Cajochen, K Kräuchi and A Wirz-Justice. ‘Role of melatonin in the regulation of human circadian rhythms and sleep’. In: *Journal of neuroendocrinology* 15.4 (2003), pp. 432–437.

[84] Josephine Arendt. ‘Melatonin and human rhythms’. In: *Chronobiology international* 23.1-2 (2006), pp. 21–37.

[85] Harris R Lieberman, Franz Waldhauser, Gail Garfield, Harry J Lynch and Richard J Wurtman. ‘Effects of melatonin on human mood and performance’. In: *Brain research* 323.2 (1984), pp. 201–207.

[86] Carsten T Beuckmann and Masashi Yanagisawa. ‘Orexins: from neuropeptides to energy homeostasis and sleep/wake regulation’. In: *Journal of molecular medicine* 80.6 (2002), pp. 329–342.

[87] Nava Zisapel. ‘Melatonin–dopamine interactions: from basic neurochemistry to a clinical setting’. In: *Cellular and molecular neurobiology* 21.6 (2001), pp. 605–616.

[88] Lotta Arborelius, MJ Owens, PM Plotsky and Charles B Nemeroff. ‘The role of corticotropin-releasing factor in depression and anxiety disorders.’ In: *The Journal of endocrinology* 160.1 (1999), pp. 1–12.

[89] B Claustrat and J Leston. ‘Melatonin: Physiological effects in humans’. In: *Neurochirurgie* 61.2-3 (2015), pp. 77–84.

[90] Takeshi Sakurai et al. ‘Input of orexin/hypocretin neurons revealed by a genetically encoded tracer in mice’. In: *Neuron* 46.2 (2005), pp. 297–308.

[91] Kyoko Yoshida, Sarah McCormack, Rodrigo A España, Amanda Crocker and Thomas E Scammell. ‘Afferents to the orexin neurons of the rat brain’. In: *Journal of Comparative Neurology* 494.5 (2006), pp. 845–861.

[92] Takeshi Sakurai and Michihiro Mieda. ‘Connectomics of orexin-producing neurons: interface of systems of emotion, energy homeostasis and arousal’. In: *Trends in pharmacological sciences* 32.8 (2011), pp. 451–462.

[93] Natsuko Tsujino and Takeshi Sakurai. ‘Orexin/hypocretin: a neuropeptide at the interface of sleep, energy homeostasis, and reward system’. In: *Pharmacological reviews* 61.2 (2009), pp. 162–176.

[94] Takeshi Sakurai. ‘Roles of orexins and orexin receptors in central regulation of feeding behavior and energy homeostasis’. In: *CNS & Neurological Disorders-Drug Targets (Formerly Current Drug Targets-CNS & Neurological Disorders)* 5.3 (2006), pp. 313–325.

[95] Minoru Narita et al. ‘Direct involvement of orexinergic systems in the activation of the mesolimbic dopamine pathway and related behaviors induced by morphine’. In: *Journal of Neuroscience* 26.2 (2006), pp. 398–405.

[96] Patrice Bourgin et al. 'Hypocretin-1 modulates rapid eye movement sleep through activation of locus coeruleus neurons'. In: *Journal of neuroscience* 20.20 (2000), pp. 7760–7765.

[97] Akihiro Yamanaka et al. 'Hypothalamic orexin neurons regulate arousal according to energy balance in mice'. In: *Neuron* 38.5 (2003), pp. 701–713.

[98] Kun-Ruey Shieh, Yeh-Shiu Chu and Jenn-Tser Pan. 'Circadian change of dopaminergic neuron activity: effects of constant light and melatonin'. In: *Neuroreport* 8.9 (1997), pp. 2283–2287.

[99] Jeffrey H Meyer et al. 'Elevated putamen D 2 receptor binding potential in major depression with motor retardation: an [11 C] raclopride positron emission tomography study'. In: *American Journal of Psychiatry* 163.9 (2006), pp. 1594–1602.

[100] Kent C Berridge. 'The debate over dopamine's role in reward: the case for incentive salience'. In: *Psychopharmacology* 191.3 (2007), pp. 391–431.

[101] Marco Leyton. 'The neurobiology of desire: Dopamine and the regulation of mood and motivational states in humans.' In: *Pleasures of the Brain*. Ed. by Berridge KC Kringelbach ML. New York: Oxford University Press, 2010.

[102] Rodrigo A Bressan and Jose A Crippa. 'The role of dopamine in reward and pleasure behaviour–review of data from preclinical research'. In: *Acta Psychiatrica Scandinavica* 111 (2005), pp. 14–21.

[103] Jeongah Kim, Sangwon Jang, Han Kyoung Choe, Sooyoung Chung, Gi Hoon Son and Kyungjin Kim. 'Implications of circadian rhythm in dopamine and mood regulation'. In: *Molecules and cells* 40.7 (2017), p. 450.

[104] L Voruganti and AG Awad. 'Neuroleptic dysphoria: towards a new synthesis'. In: *Psychopharmacology* 171.2 (2004), pp. 121–132.

[105] Karim Fifel and Howard M Cooper. 'Loss of dopamine disrupts circadian rhythms in a mouse model of Parkinson's disease'. In: *Neurobiology of disease* 71 (2014), pp. 359–369.

[106] Dusana Dorjee. *Mind, brain and the path to happiness: A guide to Buddhist mind training and the neuroscience of meditation*. Routledge, 2013.

[107] Elena Baixauli Gallego. 'Happiness: Role of Dopamine and Serotonin on mood and negative emotions'. In: *Emergency Medicine* 6.2 (2017), pp. 33–51.

[108] EJ Marijke Achterberg et al. 'Contrasting roles of dopamine and noradrenaline in the motivational properties of social play behavior in rats'. In: *Neuropsychopharmacology* 41.3 (2016), pp. 858–868.

[109] Catherine Jonnakuty and Claudia Gragnoli. 'What do we know about serotonin?' In: *Journal of cellular physiology* 217.2 (2008), pp. 301–306.

[110] Rachel LC Mitchell and Louise H Phillips. ‘The psychological, neurochemical and functional neuroanatomical mediators of the effects of positive and negative mood on executive functions’. In: *Neuropsychologia* 45.4 (2007), pp. 617–629.

[111] Thomas F Giustino, Paul J Fitzgerald, Reed L Ressler and Stephen Maren. ‘Locus coeruleus toggles reciprocal prefrontal firing to reinstate fear’. In: *Proceedings of the National Academy of Sciences* 116.17 (2019), pp. 8570–8575.

[112] Siddhartha Joshi, Yin Li, Rishi M Kalwani and Joshua I Gold. ‘Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex’. In: *Neuron* 89.1 (2016), pp. 221–234.

[113] Jane M Thompson, Chris J O’Callaghan, Bronwyn A Kingwell, Gavin W Lambert, Garry L Jennings and Murray D Esler. ‘Total norepinephrine spillover, muscle sympathetic nerve activity and heart-rate spectra analysis in a patient with dopamine β-hydroxylase deficiency’. In: *Journal of the autonomic nervous system* 55.3 (1995), pp. 198–206.

[114] Lane K Bekar, Helen S Wei and Maiken Nedergaard. ‘The locus coeruleus-norepinephrine network optimizes coupling of cerebral blood volume with oxygen demand’. In: *Journal of Cerebral Blood Flow & Metabolism* 32.12 (2012), pp. 2135–2145.

[115] PD Lambert et al. ‘A role for neuropeptide-Y, dynorphin, and noradrenaline in the central control of food intake after food deprivation’. In: *Endocrinology* 133.1 (1993), pp. 29–32.

[116] SL Foote, G Aston-Jones and FE Bloom. ‘Impulse activity of locus coeruleus neurons in awake rats and monkeys is a function of sensory stimulation and arousal’. In: *Proceedings of the National Academy of Sciences* 77.5 (1980), pp. 3033–3037.

[117] Ebony R Samuels and Elemer Szabadi. ‘Functional neuroanatomy of the noradrenergic locus coeruleus: its roles in the regulation of arousal and autonomic function part I: principles of functional organisation’. In: *Current neuropharmacology* 6.3 (2008), pp. 235–253.

[118] George F Koob. ‘Corticotropin-releasing factor, norepinephrine, and stress’. In: *Biological psychiatry* 46.9 (1999), pp. 1167–1180.

[119] David A Morilak et al. ‘Role of brain norepinephrine in the behavioral response to stress’. In: *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 29.8 (2005), pp. 1214–1224.

[120] Tania Singer et al. ‘Effects of oxytocin and prosocial behavior on brain responses to direct and vicariously experienced pain.’ In: *Emotion* 8.6 (2008), p. 781.

[121] Markus Heinrichs and Gregor Domes. ‘Neuropeptides and social behaviour: effects of oxytocin and vasopressin in humans’. In: *Progress in brain research* 170 (2008), pp. 337–350.

[122] Markus Heinrichs, Bernadette von Dawans and Gregor Domes. 'Oxytocin, vasopressin, and human social behavior'. In: *Frontiers in neuroendocrinology* 30.4 (2009), pp. 548–557.

[123] Inga D Neumann, A Wigger, L Torner, F Holsboer and R Landgraf. 'Brain oxytocin inhibits basal and stress-induced activity of the hypothalamo-pituitary-adrenal axis in male and female rats: partial action within the paraventricular nucleus.' In: *Journal of neuroendocrinology* (2000).

[124] Markus Heinrichs, Thomas Baumgartner, Clemens Kirschbaum and Ulrike Ehlert. 'Social support and oxytocin interact to suppress cortisol and subjective responses to psychosocial stress'. In: *Biological psychiatry* 54.12 (2003), pp. 1389–1398.

[125] Sara B Algoe, Laura E Kurtz and Karen Grewen. 'Oxytocin and social bonds: The role of oxytocin in perceptions of romantic partners' bonding behavior'. In: *Psychological science* 28.12 (2017), pp. 1763–1772.

[126] Anne Campbell. 'Oxytocin and human social behavior'. In: *Personality and Social Psychology Review* 14.3 (2010), pp. 281–295.

[127] Michael Kosfeld, Markus Heinrichs, Paul J Zak, Urs Fischbacher and Ernst Fehr. 'Oxytocin increases trust in humans'. In: *Nature* 435.7042 (2005), pp. 673–676.

[128] Jorgea Barraza and Paulj Zak. 'Empathy toward strangers triggers oxytocin release and subsequent generosity'. In: *Annals of the New York Academy of Sciences* 1167.1 (2009), pp. 182–189.

[129] Carolyn H Declerck, Christophe Boone and Toko Kiyonari. 'Oxytocin and cooperation under conditions of uncertainty: the modulating role of incentives and social information'. In: *Hormones and behavior* 57.3 (2010), pp. 368–374.

[130] A Kalsbeek, R Van der Spek, J Lei, E Endert, RM Buijs and E Fliers. 'Circadian rhythms in the hypothalamo–pituitary–adrenal (HPA) axis'. In: *Molecular and cellular endocrinology* 349.1 (2012), pp. 20–29.

[131] Oliver J Bosch and Inga D Neumann. 'Vasopressin released within the central amygdala promotes maternal aggression'. In: *European Journal of Neuroscience* 31.5 (2010), pp. 883–891.

[132] Eléonore Beurel and Charles B Nemeroff. 'Interaction of stress, corticotropin-releasing factor, arginine vasopressin and behaviour'. In: *Behavioral Neurobiology of Stress-related Disorders* (2014), pp. 67–80.

[133] Jia-Da Li, Katherine J Burton, Chengkang Zhang, Shuang-Bao Hu and Qun-Yong Zhou. 'Vasopressin receptor V1a regulates circadian rhythms of locomotor activity and expression of clock-controlled genes in the suprachiasmatic nuclei'. In: *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 296.3 (2009), R824–R830.

[134] A Kalsbeek, E Fliers, MA Hofman, DF Swaab and RM Buijs. ‘Vasopressin and the output of the hypothalamic biological clock’. In: *Journal of neuroendocrinology* 22.5 (2010), pp. 362–372.

[135] Inga D Neumann and Rainer Landgraf. ‘Balance of brain oxytocin and vasopressin: implications for anxiety, depression, and social behaviors’. In: *Trends in neurosciences* 35.11 (2012), pp. 649–659.

[136] Nuria De Pedro, Angel Luis Alonso-Gómez, Begoña Gancedo, María Jesús Delgado and Mercedes Alonso-Bedate. ‘Role of corticotropin-releasing factor (CRF) as a food intake regulator in goldfish’. In: *Physiology & behavior* 53.3 (1993), pp. 517–520.

[137] Azra Jaferi and Seema Bhatnagar. ‘Corticotropin-releasing hormone receptors in the medial prefrontal cortex regulate hypothalamic–pituitary–adrenal activity and anxiety-related behavior regardless of prior stress experience’. In: *Brain research* 1186 (2007), pp. 212–223.

[138] Tatyana Strekalova, Rainer Spanagel, Dusan Bartsch, Fritz A Henn and Peter Gass. ‘Stress-induced anhedonia in mice is associated with deficits in forced swimming and exploration’. In: *Neuropsychopharmacology* 29.11 (2004), pp. 2007–2017.

[139] Karel Pacak, Miklos Palkovits, Irwin J Kopin and David S Goldstein. ‘Stress-induced norepinephrine release in the hypothalamic paraventricular nucleus and pituitary-adrenocortical and sympathoadrenal activity: in vivo microdialysis studies’. In: *Frontiers in neuroendocrinology* 16.2 (1995), pp. 89–150.

[140] Daniel M Keenan, Ferdinand Roelfsema and Johannes D Veldhuis. ‘Endogenous ACTH concentration-dependent drive of pulsatile cortisol secretion in the human’. In: *American Journal of Physiology-Endocrinology and Metabolism* 287.4 (2004), E652–E661.

[141] Mary F Dallman et al. ‘Feast and famine: critical role of glucocorticoids with insulin in daily energy flow’. In: *Frontiers in neuroendocrinology* 14.4 (1993), pp. 303–347.

[142] Jay Schulkin et al. *The neuroendocrine regulation of behavior*. Cambridge University Press, 1999.

[143] Sarah Leach and Kazuhiro Suzuki. ‘Adrenergic Signaling in Circadian Control of Immunity’. In: *Frontiers in Immunology* 11 (2020), p. 1235.

[144] Frank AJL Scheer et al. ‘Impact of the human circadian system, exercise, and their interaction on cardiovascular function’. In: *Proceedings of the National Academy of Sciences* 107.47 (2010), pp. 20541–20546.

[145] Anthony JM Verberne, Willian S Korim, Azadeh Sabetghadam and Ida J Llewellyn-Smith. ‘Adrenaline: insights into its metabolic roles in hypoglycaemia and diabetes’. In: *British journal of pharmacology* 173.9 (2016), pp. 1425–1437.

[146] Gerhard Stemmler, Tatjana Aue and Jan Wacker. 'Anger and fear: Separable effects of emotion and motivational direction on somatovisceral responses'. In: *International journal of psychophysiology* 66.2 (2007), pp. 141–153.

[147] Terry McMorris et al. 'Heat stress, plasma concentrations of adrenaline, noradrenaline, 5-hydroxytryptamine and cortisol, mood state and cognitive performance'. In: *International Journal of Psychophysiology* 61.2 (2006), pp. 204–215.

[148] Derk-Jan Dijk and Malcolm von Schantz. 'Timing and consolidation of human sleep, wakefulness, and performance by a symphony of oscillators'. In: *Journal of biological rhythms* 20.4 (2005), pp. 279–290.

[149] Edmund T Rolls. 'Limbic systems for emotion and for memory, but no single limbic system'. In: *Cortex* 62 (2015), pp. 119–157.

[150] Zoe R Donaldson and Larry J Young. 'Oxytocin, vasopressin, and the neurogenetics of sociality'. In: *Science* 322.5903 (2008), pp. 900–904.

[151] EW Lamont and S Amir. 'Circadian and ultradian clocks/rhythms'. In: *Encyclopedia of Behavioral Neuroscience*. Elsevier, 2010.

[152] Urs Albrecht. 'Circadian clocks and mood-related behaviors'. In: *Circadian clocks* (2013), pp. 227–239.

[153] Charna Dibner, Ueli Schibler and Urs Albrecht. 'The mammalian circadian timing system: organization and coordination of central and peripheral clocks'. In: *Annual review of physiology* 72 (2010), pp. 517–549.

[154] Kwangwook Cho, A Ennaceur, Jon C Cole and Chang Kook Suh. 'Chronic jet lag produces cognitive deficits'. In: *Journal of Neuroscience* 20.6 (2000), RC66–RC66.

[155] Robert Y Moore and Nicholas J Lenn. 'A retinohypothalamic projection in the rat'. In: *Journal of Comparative Neurology* 146.1 (1972), pp. 1–14.

[156] David L Felten and John R Sladek Jr. 'Monoamine distribution in primate brain V. Monoaminergic nuclei: anatomy, pathways and local organization'. In: *Brain research bulletin* 10.2 (1983), pp. 171–284.

[157] Henricus G Ruhé, Nada S Mason and Aart H Schene. 'Mood is indirectly related to serotonin, norepinephrine and dopamine levels in humans: a meta-analysis of monoamine depletion studies'. In: *Molecular psychiatry* 12.4 (2007), pp. 331–359.

[158] Kirill S Korshunov, Laura J Blakemore and Paul Q Trombley. 'Dopamine: a modulator of circadian rhythms in the central nervous system'. In: *Frontiers in Cellular Neuroscience* 11 (2017), p. 91.

[159] Nora Catherine Nickels. 'The Relationship between Acute Psychosocial Stress, Hormones, Cognition, and Social Behavior in Men and Women'. In: *Knowledge Chicago University* (2019).

[160] Bruce S McEwen. ‘Stress and hippocampal plasticity’. In: *Annual review of neuroscience* 22.1 (1999), pp. 105–122.

[161] Christine Heim, D Jeffrey Newport, Tanja Mletzko, Andrew H Miller and Charles B Nemeroff. ‘The link between childhood trauma and depression: insights from HPA axis studies in humans’. In: *Psychoneuroendocrinology* 33.6 (2008), pp. 693–710.

[162] Henrik Oster et al. ‘The circadian rhythm of glucocorticoids is regulated by a gating mechanism residing in the adrenal cortical clock’. In: *Cell metabolism* 4.2 (2006), pp. 163–173.

[163] Peretz Lavie. ‘Melatonin: role in gating nocturnal rise in sleep propensity’. In: *Journal of biological rhythms* 12.6 (1997), pp. 657–665.

[164] SR Pandi-Perumal, N Zisapel, V Srinivasan and DP Cardinali. ‘Melatonin and sleep in aging population’. In: *Experimental gerontology* 40.12 (2005), pp. 911–925.

[165] Seithikurippu R Pandi-Perumal et al. ‘Physiological effects of melatonin: role of melatonin receptors and signal transduction pathways’. In: *Progress in neurobiology* 85.3 (2008), pp. 335–353.

[166] Takeshi Sakurai. ‘The neural circuit of orexin (hypocretin): maintaining sleep and wakefulness’. In: *Nature Reviews Neuroscience* 8.3 (2007), pp. 171–181.

[167] Isabella M Hower, Sara A Harper and Thomas W Buford. ‘Circadian rhythms, exercise, and cardiovascular health’. In: *Journal of circadian rhythms* 16 (2018).

[168] Martha U Gillette and Shelley A Tischkau. ‘Suprachiasmatic nucleus: the brain’s circadian clock’. In: *Recent progress in hormone research* 54.1 (1999), pp. 33–58.

[169] Philip Carew Withers, Christine E Cooper, Shane K Maloney, Francisco Bozinovic and Ariovaldo P Cruz-Neto. *Ecological and environmental physiology of mammals*. Vol. 5. Oxford University Press, 2016.

[170] John T Cacioppo and Jean Decety. ‘What are the brain mechanisms on which psychological processes are based?’ In: *Perspectives on psychological science* 4.1 (2009), pp. 10–18.

[171] James R Anderson. ‘Social stimuli and social rewards in primate learning and cognition’. In: *Behavioural processes* 42.2-3 (1998), pp. 159–175.

[172] Dan Sperber and Lawrence Hirschfeld. ‘Culture, cognition, and evolution’. In: *MIT encyclopedia of the cognitive sciences* (1999), pp. 111–132.

[173] Valeria Rubino et al. ‘Activity in medial prefrontal cortex during cognitive evaluation of threatening stimuli as a function of personality style’. In: *Brain research bulletin* 74.4 (2007), pp. 250–257.

[174] Michael A Arbib and Jean-Marc Fellous. ‘Emotions: from brain to robot’. In: *Trends in cognitive sciences* 8.12 (2004), pp. 554–561.

[175] Robert Plutchik. ‘A general psychoevolutionary theory of emotion’. In: *Theories of emotion*. Elsevier, 1980, pp. 3–33.

[176] Paul Ekman. ‘Emotions revealed’. In: *Bmj* 328.Suppl S5 (2004).

[177] Eliza Bliss-Moreau and Peter H Rudebeck. ‘Animal Models of Human Mood’. In: *Neuroscience & Biobehavioral Reviews* (2020).

[178] Marja Kokkonen and Lea Pulkkinen. ‘Examination of the paths between personality, current mood, its evaluation, and emotion regulation’. In: *European Journal of Personality* 15.2 (2001), pp. 83–104.

[179] Ricardo Santos, Goreti Marreiros, Carlos Ramos, José Neves and José Bulas-Cruz. ‘Personality, emotion, and mood in agent-based group decision making’. In: *IEEE Annals of the History of Computing* 26.06 (2011), pp. 58–66.

[180] Juan David Velásquez. ‘Cathexis–a computational model for the generation of emotions and their influence in the behavior of autonomous agents’. PhD thesis. Massachusetts Institute of Technology, 1996.

[181] Carroll E Izard. ‘Four systems for emotion activation: Cognitive and noncognitive processes.’ In: *Psychological review* 100.1 (1993), p. 68.

[182] Silvan S Tomkins. *Affect imagery consciousness: Volume I: The positive affects*. Springer publishing company, 1962.

[183] Andrew Ortony, Gerald L Clore and Allan Collins. *The cognitive structure of emotions*. Cambridge university press, 1990.

[184] Aaron Sloman. ‘Prolegomena to a theory of communication and affect’. In: *Communication from an artificial intelligence perspective*. Springer, 1992, pp. 229–260.

[185] Paul Ekman. ‘Basic emotions’. In: *Handbook of cognition and emotion* 98.45-60 (1999), p. 16.

[186] Edmund T Rolls. ‘Precis of the brain and emotion’. In: *Behavioral and brain sciences* 23.2 (2000), pp. 177–191.

[187] Nico H Frijda. *The laws of emotion*. Psychology Press, 2017.

[188] Nico H Frijda and Jaap Swagerman. ‘Can computers feel? Theory and design of an emotional system’. In: *Cognition and emotion* 1.3 (1987), pp. 235–257.

[189] Nico H Frijda. ‘Emotions in robots’. In: *Compatible approaches to cognitive science* (1995), pp. 501–516.

[190] James A Russell. ‘A circumplex model of affect.’ In: *Journal of personality and social psychology* 39.6 (1980), p. 1161.

[191] Hugo Lövheim. ‘A new three-dimensional model for emotions and monoamine neurotransmitters’. In: *Medical hypotheses* 78.2 (2012), pp. 341–348.

[192] Silvan S Tomkins. ‘Affect theory’. In: *Approaches to emotion* 163.163-195 (1984).

[193] Alexey Leukhin, Max Talanov, Jordi Vallverdú and Fail Gafarov. ‘Bio-plausible simulation of three monoamine systems to replicate emotional phenomena in a machine’. In: *Biologically inspired cognitive architectures* 26 (2018), pp. 166–173.

[194] Jordi Vallverdú, Max Talanov, Salvatore Distefano, Manuel Mazzara, Alexander Tchitchigin and Ildar Nurgaliev. ‘A cognitive architecture for the implementation of emotions in computing systems’. In: *Biologically Inspired Cognitive Architectures* 15 (2016), pp. 34–40.

[195] Juzheng Zhang, Jianmin Zheng and Nadia Magnenat-Thalmann. ‘Modeling personality, mood, and emotions’. In: *Context aware human-robot and human-agent interaction*. Springer, 2016, pp. 211–236.

[196] Xianyu Qi, Wei Wang, Lei Guo, Mingbo Li, Xiaoyu Zhang and Ran Wei. ‘Building a plutchik’s wheel inspired affective model for social robots’. In: *Journal of Bionic Engineering* 16.2 (2019), pp. 209–221.

[197] Konrad Lorenz. *Motivation of human and animal behavior*. Litton Educational Publishing, 1973.

[198] Ronald C Arkin, Masahiro Fujita, Tsuyoshi Takagi and Rika Hasegawa. ‘An ethological and emotional basis for human–robot interaction’. In: *Robotics and Autonomous Systems* 42.3-4 (2003), pp. 191–201.

[199] María Malfaz, Álvaro Castro-González, Ramón Barber and Miguel A Salichs. ‘A biologically inspired architecture for an autonomous and social robot’. In: *IEEE Transactions on Autonomous Mental Development* 3.3 (2011), pp. 232–246.

[200] Marcos Maroto-Gómez, Álvaro Castro-González, José Carlos Castillo, María Malfaz and Miguel A Salichs. ‘A bio-inspired motivational decision making system for social robots based on the perception of the user’. In: *Sensors* 18.8 (2018), p. 2691.

[201] Abraham Harold Maslow. ‘A theory of human motivation.’ In: *Psychological review* 50.4 (1943), p. 370.

[202] Cade McCall and Tania Singer. ‘The animal and human neuroendocrinology of social cognition, motivation and behavior’. In: *Nature neuroscience* 15.5 (2012), pp. 681–688.

[203] JM Koolhaas, SF De Boer and B Bohus. ‘Motivational systems or motivational states: behavioural and physiological evidence’. In: *Applied Animal Behaviour Science* 53.1-2 (1997), pp. 131–143.

[204] Sebastian Thrun et al. ‘MINERVA: A second-generation museum tour-guide robot’. In: *Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No. 99CH36288C)*. Vol. 3. IEEE. 1999.

[205] Wolfram Burgard et al. ‘The museum tour-guide robot RHINO’. In: *Autonome Mobile Systeme 1998*. Springer, 1999, pp. 245–254.

[206] Cynthia Breazeal. 'Toward sociable robots'. In: *Robotics and autonomous systems* 42.3-4 (2003), pp. 167–175.

[207] Cynthia Breazeal et al. 'A motivational system for regulating human-robot interaction'. In: *Association for the Advancement of Artificial Intelligence*. 1998, pp. 54–61.

[208] Junichi Osada, Shinichi Ohnaka and Miki Sato. 'The scenario and design process of childcare robot, PaPeRo'. In: *Proceedings of the 2006 ACM SIGCHI international conference on Advances in computer entertainment technology*. 2006, 80–es.

[209] Albert van Breemen, Xue Yan and Bernt Meerbeek. 'iCat: an animated user-interface robot with personality'. In: *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*. 2005, pp. 143–144.

[210] Víctor Gonzalez-Pacheco, Arnaud Ramey, Fernando Alonso-Martín, A Castro-González and Miguel A Salichs. 'Maggie: A social robot as a gaming platform'. In: *International Journal of Social Robotics* 3.4 (2011), pp. 371–381.

[211] F Alonso-Martin, Arnaud A Ramey and Miguel A Salichs. 'Maggie: el robot traductor'. In: *Proceedings of the 9th Workshop RoboCity2030-II, Madrid, Spain*. 2011, pp. 57–73.

[212] J Salichs, Á Castro-González and Miguel Ángel Salichs. 'Infrared remote control with a social robot'. In: *FIRA RoboWorld congress*. Springer. 2009, pp. 86–95.

[213] Diego Rodriguez-Losada, Fernando Matia, Ramon Galan, Miguel Hernando, Juan Manuel Montero and Juan Manuel Lucas. *Urbano, an interactive mobile tour-guide robot*. IntechOpen, 2008.

[214] J Javier Rainer, Salvador Cobos-Guzman and Ramón Galán. 'Decision making algorithm for an autonomous guide-robot using fuzzy logic'. In: *Journal of Ambient Intelligence and Humanized Computing* 9.4 (2018), pp. 1177–1189.

[215] Ross Mead, Mark Yim, MJ Matarić, Simon Kim and Braden McDorman. 'Quori: A community-driven modular social robot platform for human-robot interaction'. In: *2015 International Conference on Social Robotics (ICSR 2015) Robot Design Competition*. 2015.

[216] Arlene Mannion et al. 'Introducing the social robot MARIO to people living with dementia in long term residential care: Reflections'. In: *International Journal of Social Robotics* 12.2 (2020), pp. 535–547.

[217] Luigi Asprino, Paolo Ciancarini, Andrea Giovanni Nuzzolese, Valentina Presutti and Alessandro Russo. 'A reference architecture for social robots'. In: *Journal of Web Semantics* (2021), p. 100683.

[218] Ruby Yu et al. 'Use of a therapeutic, socially assistive pet robot (PARO) in improving mood and stimulating social interaction and communication for people with dementia: Study protocol for a randomized controlled trial'. In: *JMIR research protocols* 4.2 (2015), e4189.

[219] Wendy Moyle et al. 'Use of a robotic seal as a therapeutic tool to improve dementia symptoms: a cluster-randomized controlled trial'. In: *Journal of the American Medical Directors Association* 18.9 (2017), pp. 766–773.

[220] In Soon Koh and Hee Sun Kang. 'Effects of intervention using PARO on the cognition, emotion, problem behavior, and social interaction of elderly people with dementia'. In: *Journal of Korean Academy of Community Health Nursing* 29.3 (2018), pp. 300–309.

[221] Cindy Jones et al. 'Does cognitive impairment and agitation in dementia influence intervention effectiveness? Findings from a cluster-randomized-controlled trial with the therapeutic robot, PARO'. In: *Journal of the American Medical Directors Association* 19.7 (2018), pp. 623–626.

[222] Patricia A Kelly et al. 'The effect of PARO robotic seals for hospitalized patients with dementia: a feasibility study'. In: *Geriatric Nursing* 42.1 (2021), pp. 37–45.

[223] Masahiro Fujita. 'AIBO: Toward the era of digital creatures'. In: *The International Journal of Robotics Research* 20.10 (2001), pp. 781–794.

[224] Ronald C Arkin. 'Moving up the food chain: Motivation and Emotion in behavior-based robots'. In: *Who needs emotions?: The brain meets the robot*. Georgia Institute of Technology, 2003, pp. 245–269.

[225] Ylva Fernaeus, Maria Håkansson, Mattias Jacobsson and Sara Ljungblad. 'How do you play with a robotic toy animal? A long-term study of Pleo'. In: *Proceedings of the 9th international Conference on interaction Design and Children*. 2010, pp. 39–48.

[226] Karola Pitsch and Benjamin Koch. 'How infants perceive the toy robot pleo. an exploratory case study on infant-robot-interaction'. In: *Second International Symposium on New Frontiers in Human-Robot-Interaction (AISB), SSAISB: The Society for the Study of Artificial Intelligence and the Simulation of Behavior, Leicester, UK*. Citeseer. 2010.

[227] Adso Fernández-Baena, Roger Boldú, Jordi Albo-Canals and David Miralles. 'Interaction between Vleo and Pleo, a virtual social character and a social robot'. In: *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE. 2015, pp. 694–699.

[228] Reensina Eind and Marcel Heerink. 'Evaluation of the use of a Pleo robot at a child consultation clinic'. In: *Proceedings of the third international conference on social robots in therapy and education. Panama*. 2018, pp. 41–43.

[229] Hideki Kozima, Marek P Michalowski and Cocoro Nakagawa. ‘Keepon’. In: *International Journal of Social Robotics* 1.1 (2009), pp. 3–18.

[230] Adam Robaczewski, Julie Bouchard, Kevin Bouchard and Sébastien Gaboury. ‘Socially assistive robots: The specific case of the NAO’. In: *International Journal of Social Robotics* (2020), pp. 1–37.

[231] Syamimi Shamsuddin et al. ‘Initial response of autistic children in human-robot interaction therapy with humanoid robot NAO’. In: *2012 IEEE 8th International Colloquium on Signal Processing and its Applications*. IEEE. 2012, pp. 188–193.

[232] Adriana Tapus et al. ‘Children with autism social engagement in interaction with Nao, an imitative robot: A series of single case experiments’. In: *Interaction studies* 13.3 (2012), pp. 315–347.

[233] Chris Lytridis, Eleni Vrochidou, Stamatis Chatzistamatis and Vassilis Kaburlasos. ‘Social engagement interaction games between children with Autism and humanoid robot NAO’. In: *The 13th international conference on soft computing models in industrial and environmental applications*. Springer. 2018, pp. 562–570.

[234] Hoang-Long Cao et al. ‘Robot-assisted joint attention: A comparative study between children with autism spectrum disorder and typically developing children in interaction with NAO’. In: *IEEE Access* 8 (2020), pp. 223325–223334.

[235] Hal Hodson. *The first family robot*. 2014.

[236] Amit Kumar Pandey and Rodolphe Gelin. ‘A mass-produced sociable humanoid robot: Pepper: The first machine of its kind’. In: *IEEE Robotics & Automation Magazine* 25.3 (2018), pp. 40–48.

[237] Iina Aaltonen, Anne Arvola, Päivi Heikkilä and Hanna Lammi. ‘Hello Pepper, may I tickle you? Children’s and adults’ responses to an entertainment robot at a shopping mall’. In: *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. 2017, pp. 53–54.

[238] Dario Allegra, Francesco Alessandro, Corrado Santoro and Filippo Stanco. ‘Experiences in using the pepper robotic platform for museum assistance applications’. In: *2018 25th IEEE International Conference on Image Processing (ICIP)*. IEEE. 2018, pp. 1033–1037.

[239] Josef Guggemos, Sabine Seufert and Stefan Sonderegger. ‘Humanoid robots in higher education: Evaluating the acceptance of Pepper in the context of an academic writing course using the UTAUT’. In: *British Journal of Educational Technology* 51.5 (2020), pp. 1864–1883.

[240] Emily C Collins, Tony J Prescott, Ben Mitchinson and Sebastian Conran. ‘MIRO: a versatile biomimetic edutainment robot’. In: *Proceedings of the 12th International Conference on Advances in Computer Entertainment Technology*. 2015, pp. 1–4.

[241] Maggie Redden. ‘Sophia: The Intersection of Artificial Intelligence and Human Rights’. In: *J. Glob. Rts. & Org.* 10 (2019), p. 155.

[242] Sara Cooper, Alessandro Di Fava, Carlos Vivas, Luca Marchionni and Francesco Ferro. ‘ARI: The social assistive robot and companion’. In: *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE. 2020, pp. 745–751.

[243] Clark Davidson Elliott. ‘The affective reasoner: a process model of emotions in a multiagent system’. PhD thesis. Northwestern University, 1992.

[244] Charles Webster. ‘Adaptive depression, affective computing, and intelligent processing’. In: *1997 IEEE International Conference on Intelligent Processing Systems (Cat. No. 97TH8335)*. Vol. 2. IEEE. 1997, pp. 1181–1184.

[245] J Velasquez. ‘Building affective robots’. In: *Proceedings of Human Robotics* (1999).

[246] Takanori Shibata and Kazuo Tanie. ‘Physical and affective interaction between human and mental commit robot’. In: *Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation (Cat. No. 01CH37164)*. Vol. 3. IEEE. 2001, pp. 2572–2577.

[247] Matthias Scheutz. ‘Affective Action Selection and Behavior Arbitration for autonomous Robots.’ In: *IC-AI*. 2002, pp. 334–340.

[248] Jean-Marc Fellous. ‘From human emotions to robot emotions’. In: *Architectures for Modeling Emotion: Cross-Disciplinary Foundations, American Association for Artificial Intelligence* (2004), pp. 39–46.

[249] Jean-Marc Fellous and Michael A Arbib. *Who needs emotions?: The brain meets the robot*. Oxford University Press, 2005.

[250] Orlando Avila-Garcia and Lola Canamero. ‘Hormonal modulation of perception in motivation-based action selection architectures’. In: *Proceedings of the Symposium on Agents that Want and Like*. The Society for the Study of Artificial Intelligence and Simulation of Behaviour. 2005.

[251] George Konidaris and Andrew Barto. ‘An adaptive robot motivational system’. In: *International Conference on Simulation of Adaptive Behavior*. Springer. 2006, pp. 346–356.

[252] Samir Alili, Rachid Alami and Vincent Montreuil. ‘A task planner for an autonomous social robot’. In: *Distributed autonomous robotic systems 8*. Springer, 2009, pp. 335–344.

[253] Christian Balkenius, Jan Morén and Stefan Winberg. ‘Interactions between motivation, emotion and attention: From biology to robotics.’ In: *IEEE International Conference on Development and Learning*. 2009.

[254] Hooman Aghaebrahimi Samani and Elham Saadatian. 'A multidisciplinary artificial intelligence model of an affective robot'. In: *International Journal of Advanced Robotic Systems* 9.1 (2012), p. 6.

[255] Elena Corina Grigore, Andre Pereira and Brian Scassellati. 'Modeling Motivational States for Adaptive Robot Companions'. In: *2015 AAAI Fall Symposium Series*. 2015.

[256] Matthew Lewis and Lola Canamero. 'Hedonic quality or reward? A study of basic pleasure in homeostasis and decision making of a motivated autonomous robot'. In: *Adaptive Behavior* 24.5 (2016), pp. 267–291.

[257] José-Antonio Cervantes, Luis-Felipe Rodríguez, Sonia López, Félix Ramos and Francisco Robles. 'Autonomous agents and ethical decision-making'. In: *Cognitive Computation* 8.2 (2016), pp. 278–296.

[258] Carole Adam, Wafa Johal, Damien Pellier, Humbert Fiorino and Sylvie Pesty. 'Social human-robot interaction: a new cognitive and affective interaction-oriented architecture'. In: *International conference on social robotics*. Springer. 2016, pp. 253–263.

[259] Sebastian Schneider, Michael Goerlich and Franz Kummert. 'A framework for designing socially assistive robot interactions'. In: *Cognitive Systems Research* 43 (2017), pp. 301–312.

[260] Zdzisław Kowalczuk and Michał Czubenko. 'An intelligent decision-making system for autonomous units based on the mind model'. In: *2018 23rd International Conference on Methods & Models in Automation & Robotics (MMAR)*. IEEE. 2018, pp. 1–6.

[261] Alejandro Romero, Francisco Bellas, Abraham Prieto and Richard J Duro. 'Utility Model Re-description within a Motivational System for Cognitive Robotics'. In: *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2018, pp. 2324–2329.

[262] Ryan J McCall, Stan Franklin, Usef Faghihi, Javier Snaider and Sean Kugele. 'Artificial Motivation for Cognitive Software Agents'. In: *Journal of Artificial General Intelligence* 11.1 (2020), pp. 38–69.

[263] Agnese Augello, Ignazio Infantino, Salvatore Gaglio, Umberto Maniscalco, Giovanni Pilato and Filippo Vella. 'An Artificial Soft Somatosensory System for a Cognitive Robot'. In: *2020 Fourth IEEE International Conference on Robotic Computing (IRC)*. IEEE. 2020, pp. 319–326.

[264] Ana Tanevska, Francesco Rea, Giulio Sandini and Alessandra Sciutti. 'Designing an affective cognitive architecture for human-humanoid interaction'. In: *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. 2018, pp. 253–254.

[265] Ana Tanevska, Francesco Rea, Giulio Sandini, Lola Cañamero and Alessandra Sciutti. 'A socially adaptable framework for human-robot interaction'. In: *Frontiers in Robotics and AI* 7 (2020).

[266] Alexander Hong et al. 'A Multimodal Emotional Human-Robot Interaction Architecture for Social Robots Engaged in Bidirectional Communication'. In: *IEEE transactions on cybernetics* (2020).

[267] José Carlos González, José Carlos Pulido and Fernando Fernández. 'A three-layer planning architecture for the autonomous control of rehabilitation therapies based on social robots'. In: *Cognitive Systems Research* 43 (2017), pp. 232–249.

[268] Dolores Cañamero. 'Modeling motivations and emotions as a basis for intelligent behavior'. In: *Proceedings of the first international conference on Autonomous agents*. 1997, pp. 148–155.

[269] Werner Plihal, Rosemarie Krug, Reinhard Pietrowsky, Horst L Fehm and Jan Born. 'Corticosteroid receptor mediated effects on mood in humans'. In: *Psychoneuroendocrinology* 21.6 (1996), pp. 515–523.

[270] Thorsten Schodde, Kirsten Bergmann and Stefan Kopp. 'Adaptive robot language tutoring based on Bayesian knowledge tracing and predictive decision-making'. In: *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. 2017, pp. 128–136.

[271] Hannes Ritschel and Elisabeth André. 'Real-time robot personality adaptation based on reinforcement learning and social signals'. In: *Proceedings of the companion of the 2017 acm/ieee international conference on human-robot interaction*. 2017, pp. 265–266.

[272] Klaus Weber, Hannes Ritschel, Florian Lingenfelser and Elisabeth André. 'Real-time adaptation of a robotic joke teller based on human social signals'. In: *International Conference on Autonomous Agents and Multiagent Systems* (2018).

[273] Klaus Weber, Hannes Ritschel, Ilhan Aslan, Florian Lingenfelser and Elisabeth André. 'How to shape the humor of a robot-social behavior adaptation based on reinforcement learning'. In: *Proceedings of the 20th ACM International Conference on Multimodal Interaction*. 2018, pp. 154–162.

[274] Luefeng Chen, Min Wu, Mengtian Zhou, Jinhua She, Fangyan Dong and Kaoru Hirota. 'Information-driven multirobot behavior adaptation to emotional intention in human–robot interaction'. In: *IEEE Transactions on Cognitive and Developmental Systems* 10.3 (2017), pp. 647–658.

[275] Ana Tanevska, Francesco Rea, Giulio Sandini, Lola Cañamero and Alessandra Sciutti. 'Eager to learn vs. quick to complain? how a socially adaptive robot architecture performs with different robot personalities'. In: *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*. IEEE. 2019, pp. 365–371.

[276] Antonio Andriella, Carme Torras and Guillem Alenya. 'Short-term human–robot interaction adaptability in real-world environments'. In: *International Journal of Social Robotics* 12.3 (2020), pp. 639–657.

[277] Bahar Irfan, Aditi Ramachandran, Samuel Spaulding, Dylan F Glas, Iolanda Leite and Kheng Lee Koay. 'Personalization in long-term human-robot interaction'. In: *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE. 2019, pp. 685–686.

[278] Muneeb Ahmad, Omar Mubin and Joanne Orlando. 'A systematic review of adaptivity in human-robot interaction'. In: *Multimodal Technologies and Interaction* 1.3 (2017), p. 14.

[279] Satinder P. Singh and Richard S. Sutton. 'Reinforcement Learning with Replacing Eligibility Traces'. In: *Machine Learning* 22.1 ASN - 1573-0565 (Jan. 1996), pp. 123–158.

[280] Richard S. Sutton. 'Dyna, an Integrated Architecture for Learning, Planning, and Reacting'. In: *SIGART Bulletin* 2.4 (July 1991), pp. 160–163.

[281] Hüllermeier E. Fürnkranz J. 'Preference Learning: An Introduction'. In: *Preference Learning*. Springer, 2010, pp. 1–17.

[282] Eyke Hüllermeier, Johannes Fürnkranz, Weiwei Cheng and Klaus Brinker. 'Label ranking by learning pairwise preferences'. In: *Artificial Intelligence* 172.16-17 (2008), pp. 1897–1916.

[283] Shankar Vembu and Thomas Gärtner. 'Label ranking algorithms: A survey'. In: *Preference learning*. Springer, 2010, pp. 45–64.

[284] Johannes Fürnkranz and Eyke Hüllermeier. 'Pairwise preference learning and ranking'. In: *European conference on machine learning*. Springer. 2003, pp. 145–156.

[285] Johannes Fürnkranz and Eyke Hüllermeier. 'Preference learning and ranking by pairwise comparison'. In: *Preference learning*. Springer, 2010, pp. 65–82.

[286] Eyke Hüllermeier and Johannes Fürnkranz. 'On predictive accuracy and risk minimization in pairwise label ranking'. In: *Journal of Computer and System Sciences* 76.1 (2010), pp. 49–62.

[287] Dorsa Sadigh, Anca D Dragan, Shankar Sastry and Sanjit A Seshia. 'Active preference-based learning of reward functions'. In: *University of California* (2017).

[288] Leo Woiceshyn, Yuchi Wang, Goldie Nejat and Beno Benhabib. 'Personalized clothing recommendation by a social robot'. In: *2017 IEEE International Symposium on Robotics and Intelligent Sensors (IRIS)*. IEEE. 2017, pp. 179–185.

[289] Yangming Zhou and Guoping Qiu. 'Random forest for label ranking'. In: *Expert Systems with Applications* 112 (2018), pp. 99–109.

[290] Cláudio Rebelo de Sá, Carlos Soares, Arno Knobbe and Paulo Cortez. 'Label ranking forests'. In: *Expert systems* 34.1 (2017), pp. 1–8.

[291] Klaus Brinker and Eyke Hüllermeier. 'Case-based label ranking'. In: *European Conference on Machine Learning*. Springer. 2006, pp. 566–573.

[292] Havi Werbin-Ofir, Lihi Dery and Erez Shmueli. 'Beyond majority: Label ranking ensembles based on voting rules'. In: *Expert Systems with Applications* 136 (2019), pp. 50–61.

[293] Miguel A Salichs et al. 'Mini: a new social robot for the elderly'. In: *International Journal of Social Robotics* 12.6 (2020), pp. 1231–1249.

[294] Fernando Alonso-Martín, Juan José Gamboa-Montero, José Carlos Castillo, Álvaro Castro-González and Miguel Ángel Salichs. 'Detecting and classifying human touches in a social robot through acoustic sensing and machine learning'. In: *Sensors* 17.5 (2017), p. 1138.

[295] Carlos M Gómez et al. 'Desarrollo de una versión de bajo coste del robot social Mini'. In: *XL Jornadas de Automática*. Universidade da Coruña, Servizo de Publicacións. 2019, pp. 718–725.

[296] Janelle Nimer and Brad Lundahl. 'Animal-assisted therapy: A meta-analysis'. In: *Anthrozoös* 20.3 (2007), pp. 225–238.

[297] Aubrey H Fine. *Handbook on animal-assisted therapy: Theoretical foundations and guidelines for practice*. academic press, 2010.

[298] Kimberly Eaton Hoagwood, Mary Acri, Meghan Morrissey and Robin Peth-Pierce. 'Animal-assisted therapies for youth with or at risk for mental health problems: A systematic review'. In: *Applied developmental science* 21.1 (2017), pp. 1–13.

[299] Joost Broekens, Marcel Heerink, Henk Rosendal et al. 'Assistive social robots in elderly care: a review'. In: *Gerontechnology* 8.2 (2009), pp. 94–103.

[300] Enrique Fernández Rodicio. 'Human-Robot Interaction architecture for interactive and lively social robots'. PhD thesis. University Carlos III of Madrid, 2021.

[301] Enrique Fernández-Rodicio, Álvaro Castro-González, Fernando Alonso-Martín, Marcos Maroto-Gómez and Miguel Á Salichs. 'Modelling Multimodal Dialogues for Social Robots Using Communicative Acts'. In: *Sensors* 20.12 (2020), p. 3440.

[302] Esther Salichs San José. 'Ayuda a personas mayores con un robot social: estimulación cognitiva'. In: *University Carlos III of Madrid* (2019).

[303] R Toni. 'The neuroendocrine system: organization and homeostatic role.' In: *Journal of endocrinological investigation* 27.6 Suppl (2004), pp. 35–47.

[304] Irenaus Eibl-Eibesfeldt. *Human ethology*. Routledge, 2017.

[305] Yuhei Kawano et al. ‘Circadian variations of urinary dopamine, norepinephrine, epinephrine and sodium in normotensive and hypertensive subjects’. In: *Nephron* 55.3 (1990), pp. 277–282.

[306] Abhijit Chakraborty, Wojciech Krzyzanski and William J Jusko. ‘Mathematical modeling of circadian cortisol concentrations using indirect response models: comparison of several methods’. In: *Journal of pharmacokinetics and biopharmaceutics* 27.1 (1999), pp. 23–43.

[307] Jamie M Zeitzer, Christine L Buckmaster, Karen J Parker, Craig M Hauck, David M Lyons and Emmanuel Mignot. ‘Circadian and homeostatic regulation of hypocretin in a primate model: implications for the consolidation of wakefulness’. In: *Journal of Neuroscience* 23.8 (2003), pp. 3555–3560.

[308] James R Sowers and Nicholas Vlachakis. ‘Circadian variation in plasma dopamine levels in man’. In: *Journal of endocrinological investigation* 7.4 (1984), pp. 341–345.

[309] CR Linsell, SL Lightman, PE Mullen, MJ Brown and RC Causon. ‘Circadian rhythms of epinephrine and norepinephrine in man’. In: *The Journal of Clinical Endocrinology & Metabolism* 60.6 (1985), pp. 1210–1215.

[310] Janet A Amico, Steven C Levin and Judy L Cameron. ‘Circadian rhythm of oxytocin in the cerebrospinal fluid of rhesus and cynomolgus monkeys: effects of castration and adrenalectomy and presence of a caudal-rostral gradient’. In: *Neuroendocrinology* 50.6 (1989), pp. 624–632.

[311] Jiang-Ning Zhou et al. ‘Alterations in arginine vasopressin neurons in the suprachiasmatic nucleus in depression’. In: *Archives of General Psychiatry* 58.7 (2001), pp. 655–662.

[312] Konrad Lorenz. *The foundations of ethology*. Springer Science & Business Media, 2013.

[313] FAJL Scheer et al. ‘Cardiovascular control by the biological clock’. In: *University of Amsterdam* (2002).

[314] Chin-An Wang, Talia Baird, Jeff Huang, Jonathan D Coutinho, Donald C Brien and Douglas P Munoz. ‘Arousal effects on pupil size, heart rate, and skin conductance in an emotional face task’. In: *Frontiers in neurology* 9 (2018), p. 1029.

[315] Christina M Spengler, Charles A Czeisler and Steven A Shea. ‘An endogenous circadian rhythm of respiratory control in humans’. In: *The Journal of physiology* 526.3 (2000), pp. 683–694.

[316] Tarja Saaresranta and Olli Polo. ‘Hormones and breathing’. In: *Chest* 122.6 (2002), pp. 2165–2182.

[317] W Adamczyk, M Tafil-Klawe, M Siekierka, P Zlomanczuk, P Weber and JJ Klawe. ‘Daily pattern of breathing in healthy young men’. In: *J Physiol Pharmacol* 59.Suppl 6 (2008), pp. 115–122.

[318] Giuseppe Barbato, Gianluca Ficca, Giovanni Muscettola, Mariateresa Fichele, Michele Beatrice and Franco Rinaldi. ‘Diurnal variation in spontaneous eye-blink rate’. In: *Psychiatry research* 93.2 (2000), pp. 145–151.

[319] Yuecui Kan, Haijun Duan, Xitong Chen, Xuewei Wang, Wenlong Xue and Weiping Hu. ‘Attentional blink affected by acute stress in women: The role of affective stimuli and attentional resources’. In: *Consciousness and cognition* 75 (2019), p. 102796.

[320] Clément Bourguignon and Kai-Florian Storch. ‘Control of rest: activity by a dopaminergic ultradian oscillator and the circadian clock’. In: *Frontiers in neurology* 8 (2017), p. 614.

[321] Yang Wang, Sophia E Kramer, Dorothea Wendt, Graham Naylor, Thomas Lunner and Adriana A Zekveld. ‘The pupil dilation response during speech perception in dark and light: The involvement of the parasympathetic nervous system in listening effort’. In: *Trends in Hearing* 22 (2018), p. 2331216518816603.

[322] Bernard Weiner. *Human motivation*. Psychology Press, 2013.

[323] Dolores Cañamero. ‘Designing emotions for activity selection in autonomous agents’. In: *Emotions in humans and artifacts* 115 (2003), p. 148.

[324] Álvaro Castro-González, María Malfaz and Miguel Angel Salichs. ‘An autonomous social robot in fear’. In: *IEEE Transactions on Autonomous Mental Development* 5.2 (2013), pp. 135–151.

[325] Orlando Avila-García, Lola Canamero and René te Boekhorst. ‘Analyzing the performance of “winner-take-all” and “voting-based” action selection policies within the two-resource problem’. In: *European Conference on Artificial Life*. Springer. 2003, pp. 733–742.

[326] Cynthia Breazeal. ‘Emotion and sociable humanoid robots’. In: *International Journal of Human-Computer Studies* 59.1-2 (2003), pp. 119–155.

[327] Orlando Avila-Garcia and Lola Cañamero. ‘Using hormonal feedback to modulate action selection in a competitive scenario’. In: *From Animals to Animats* 8 (2004), pp. 243–252.

[328] Ginevra Castellano, Santiago D Villalba and Antonio Camurri. ‘Recognising human emotions from body movement and gesture dynamics’. In: *International Conference on Affective Computing and Intelligent Interaction*. Springer. 2007, pp. 71–82.

[329] Kazunori Terada, Atsushi Yamauchi and Akira Ito. ‘Artificial emotion expression for a robot by dynamic color change’. In: *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. IEEE. 2012, pp. 314–321.

[330] Robert Plutchik. *Emotions and life: Perspectives from psychology, biology, and evolution.* American Psychological Association, 2003.

[331] Morgan Quigley et al. ‘ROS: an open-source Robot Operating System’. In: *ICRA workshop on open source software*. Vol. 3. 3.2. Kobe, Japan. 2009, p. 5.

[332] Andrea L Thomaz, Maya Cakmak and KB Clark. ‘Active social learning in humans and robots’. In: *Social learning theory: Phylogenetic considerations across animal, plant, and microbial taxa* (2013), pp. 113–28.

[333] Florian Schroff, Dmitry Kalenichenko and James Philbin. ‘Facenet: A unified embedding for face recognition and clustering’. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, pp. 815–823.

[334] Tiancheng Hu, Sumit Jha and Carlos Busso. ‘Temporal Head Pose Estimation From Point Cloud in Naturalistic Driving Conditions’. In: *IEEE Transactions on Intelligent Transportation Systems* (2021).

[335] John Brooke. ‘SUS: a retrospective’. In: *Journal of usability studies* 8.2 (2013), pp. 29–40.

[336] Irene Tor-Carroggio, Daniel Segura and Olga Soler-Vilageliu. ‘Usability as a premise of quality: first steps towards the validation of the System Usability Scale (SUS) into Spanish’. In: *Journal of Audiovisual Translation* 2.2 (2019), pp. 57–71.

[337] Tae Kyun Kim. ‘Understanding one-way ANOVA using conceptual figures’. In: *Korean journal of anesthesiology* 70.1 (2017), p. 22.

[338] Mary L McHugh. ‘Multiple comparison analysis testing in ANOVA’. In: *Biochemia medica* 21.3 (2011), pp. 203–209.

[339] J. Fürnkranz and E. Hüllermeier. *Preference Learning*. Springer Berlin Heidelberg, 2010.

[340] Thomas Olsson and Markus Salo. ‘Online user survey on current mobile augmented reality applications’. In: *2011 10th IEEE International Symposium on Mixed and Augmented Reality*. IEEE. 2011, pp. 75–84.

[341] Amancio Bouza and Abraham Bernstein. ‘(Partial) user preference similarity as classification-based model similarity’. In: *Semantic Web* 5.1 (2014), pp. 47–64.

[342] Maurice G Kendall. ‘The treatment of ties in ranking problems’. In: *Biometrika* (1945), pp. 239–251.

[343] Charles Spearman. ‘The proof and measurement of association between two things’. In: *The American journal of psychology* 100.3/4 (1987), pp. 441–471.

[344] Gregory W Corder and Dale I Foreman. *Nonparametric statistics for non-statisticians*. 2011.

[345] Patrick Schober, Christa Boer and Lothar A Schwarte. 'Correlation coefficients: appropriate use and interpretation'. In: *Anesthesia & Analgesia* 126.5 (2018), pp. 1763–1768.

[346] Tadayoshi Fushiki. 'Estimation of prediction error by using K-fold cross-validation'. In: *Statistics and Computing* 21.2 (2011), pp. 137–146.

[347] James Bergstra and Yoshua Bengio. 'Random search for hyper-parameter optimization.' In: *Journal of machine learning research* 13.2 (2012).

[348] Christoph Bartneck, Elizabeth Croft and Dana Kulic. 'Measuring the anthropomorphism, animacy, likeability, perceived intelligence and perceived safety of robots'. In: *University of Hertfordshire* (2008).

[349] Astrid Weiss and Christoph Bartneck. 'Meta analysis of the usage of the godspeed questionnaire series'. In: *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE. 2015, pp. 381–388.

[350] Nikhil Churamani et al. 'The impact of personalisation on human-robot interaction in learning scenarios'. In: *Proceedings of the 5th International Conference on Human Agent Interaction*. 2017, pp. 171–180.

[351] Daniel C Tozadore, Joao PH Valentini, Victor HS Rodrigues, Fernando ML Vendrameto, Rodrigo G Zavarizz and Roseli AF Romero. 'Towards adaptation and personalization in task based on human-robot interaction'. In: *2018 Latin American Robotic Symposium, 2018 Brazilian Symposium on Robotics (SBR) and 2018 Workshop on Robotics in Education (WRE)*. IEEE. 2018, pp. 383–389.

[352] Patrick E McKnight and Julius Najab. 'Mann-Whitney U Test'. In: *The Corsini encyclopedia of psychology* (2010), pp. 1–1.

[353] James T Croasmun and Lee Ostrom. 'Using likert-type scales in the social sciences.' In: *Journal of Adult Education* 40.1 (2011), pp. 19–22.

[354] Ankur Joshi, Saket Kale, Satish Chandel and D Kumar Pal. 'Likert scale: Explored and explained'. In: *British Journal of Applied Science & Technology* 7.4 (2015), p. 396.

[355] Marcos Maroto-Gómez, Rodrigo González, Álvaro Castro-González, María Malfaz and Miguel Ángel Salichs. 'Speeding-Up Action Learning in a Social Robot With Dyna-Q+: A Bioinspired Probabilistic Model Approach'. In: *IEEE Access* 9 (2021), pp. 98381–98397.

[356] Álvaro Castro-González, María Malfaz and Miguel A Salichs. 'Learning the selection of actions for an autonomous social robot by reinforcement learning based on motivations'. In: *International Journal of Social Robotics* 3.4 (2011), pp. 427–441.

[357] Maria Malfaz and Miguel A Salichs. 'Learning to deal with objects'. In: *2009 IEEE 8th International Conference on Development and Learning*. IEEE. 2009, pp. 1–6.

[358] Pao-Lu Hsu and Herbert Robbins. 'Complete convergence and the law of large numbers'. In: *Proceedings of the National Academy of Sciences of the United States of America* 33.2 (1947), p. 25.

[359] Colleen M Carpinella, Alisa B Wyman, Michael A Perez and Steven J Stroessner. 'The robotic social attributes scale (RoSAS) development and validation'. In: *Proceedings of the 2017 ACM/IEEE International Conference on human-robot interaction*. 2017, pp. 254–262.

[360] Patrick E McKight and Julius Najab. 'Kruskal-wallis test'. In: *The corsini encyclopedia of psychology* (2010), pp. 1–1.

[361] Dan Steinberg and Phillip Colla. 'CART: classification and regression trees'. In: *The top ten algorithms in data mining* 9 (2009), p. 179.

[362] J Ross Quinlan. 'Improved use of continuous attributes in C4. 5'. In: *Journal of artificial intelligence research* 4 (1996), pp. 77–90.

Appendices

Appendix A

Evaluation metrics

The following sections contain the description of the evaluation metrics and methods used in this thesis.

A.1. Kendall τ-b

Kendall [342] developed as a non-parametric measure of correlation between two rankings π and π', as represented in Equation A.1. It relies on the number of concordant (n_c) and discordant (n_d) pairs of the ranking, considering the number of ties in the predicted ranking π (t_i) and in the real ranking π' (u_j).

$$\tau = \frac{n_c - n_d}{\sqrt{(n_c + n_d + t_i) * (n_c + n_d + u_j))}} \tag{A.1}$$

A.2. Spearman ρ

The Spearman's $\boldsymbol{\rho}$ [343] rank-order correlation is a non-parametric measure of the monotonicity between the correspondence of two rankings π and π'. Unlike other correlation metrics such as Pearson's correlation coefficient, Spearman's ρ does not assume a normal distribution of the data, which makes it more suitable to our self-built dataset. As Equation A.2 represents, it depends on the number of pairs N available in the ranking and the distance $D(\pi, \pi')$ about the positioning of each label in both rankings.

$$\rho = 1 - \frac{6\sum_{i=1}^{N} D_i(\pi, \pi')^2}{N(N^2 - 1)} \tag{A.2}$$

A.3. Gini impurity

Gini impurity is a metric used by Classification and Regression Trees (CART) [361] to decide how to split a node of the tree into subnodes. It measures how often an element chosen randomly from a set would be incorrectly labelled if it was previously randomly labelled using a distribution representing the subset's labels. Like Equation A.3 shows, the Gini index of a class $\mathcal{X}$ can be computed as the squared summation of the probabilities $p(i|\mathcal{X})^2$ of selecting label i in class $\mathcal{X}$ considering all the available classes in the set C. The selected feature to produce the splitting of the tree is the one with the lowest Gini impurity value from all the available features. The Gini impurity ranges from [0, 1], where 0 means a perfect node splitting.

$$Gini\,(\mathcal{X}) = 1 - \sum_{i=1}^{C} p(i|\mathcal{X})^2 \tag{A.3}$$

A.4. Information gain

The information gain is a splitting criterium used in C4.5 [362], a decision three classifying method. The method considers the change of entropy $\mathcal{H}$ of a particular class $\mathcal{X}$ from one state to another that takes the information as given (Equation A.4).

$$\mathcal{IG}(\mathcal{X}, \mathcal{Y}) = \mathcal{H}(\mathcal{X}) - \mathcal{H}(\mathcal{X}, \mathcal{Y}) \tag{A.4}$$

Considering this definition, the metric information gain is based on the variation of entropy $\mathcal{H}$, a variable that measures the disorder of a collection of data. Mathematically, the entropy can be expressed as shown by Equation A.5,

$$\mathcal{H}(\mathcal{X}) = \sum_{i=1}^{C} -p_i \log_2 p_i \tag{A.5}$$

p_i is the probability of an element i in the dataset and C the total number of available classes.

A.5. Mean Squared Error (MSE)

In statistics, the Mean Squared Error (MSE) measures the average squared error between the real value and the estimated value of a variable. Mathematically, it depends on the size of the sample N, a vector of estimation $\widehat{Y}$, and a vector of real value Y. The MSE is a powerful metric that includes the variance and bias of the estimated input. On many occasions, it is also expressed as the Root Mean Squared Error $RMSE$.

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (\widehat{Y_i} - Y_i)^2 \tag{A.6}$$

Appendix B

Methodologies

This appendix contains essential methodologies used in this thesis, like applying Boltzmann's equation to action selection.

B.1. Boltzmann's action selection method

In RL algorithms, the action that the agent takes is usually selected using two different methods: e-greedy and the softmax distribution. e-greedy is a simple and popular way of deciding which action to execute. The method assigns a certain selection probability (p) to the action with the highest rating and the rest to selecting an action at random ($1 - p$). Although this method is easy to understand and implement, the exploration depends on the randomness of the action selection since all actions have the same probability of being selected. To overcome this issue, the softmax method balances the assignment of the selection likelihood according to the Q-value of each action and a parameter called temperature (τ). Intrinsically, the softmax method uses the Gibbs or Boltzmann distribution to produce the probabilities assigned to each action. Mathematically, the selection likelihood of action $p(a)$ is:

$$p(a) = \frac{e^{\frac{Q(a)}{\tau}}}{\sum_{b=1}^{N} e^{\frac{Q(b)}{\tau}}} \tag{B.1}$$

Where $Q(a)$ is the Q-value of action a, τ is the temperature, and N is the number of actions the robot can execute. High values of τ provoke that all actions have a similar probability, fostering exploration. As $\tau \rightarrow 0$, actions with higher Q-values increase their selection probabilities to the detriment of actions with lower Q-values.

B.2. K-fold cross-validation

In machine learning, the models developed to estimate variables require training and validation to measure the system's real performance. Cross-validation is a method used for analysing statistical data, assuring the independence of the data used for training and testing the model. There exists three different approaches for performing cross-validation: K-fold cross-validation, holdout method, and repeated random sub-sampling validation. In this work, we focus on describing the K-fold cross-validation method since it is the one used in our PL scenario.

This method consists in partitioning the original dataset into K subsamples with equal size. Then, one subsample is left aside for validating the model, while the other $K - 1$ subsets train the system. Once the validation has finished, the performance of the system is saved for further processing. The training and validation process loops K times, using a different subsample from the initial partition. Finally, the K results yielded by each cross-validation interaction are averaged to compute the system's real performance. The most important advantage of this method is the use of input data for training and validation at least once, guaranteeing that each input instance is used for validation just once.

In an actual application, the most common technique is 10-fold cross-validation, but other approaches like maintaining an unfixed K value that increases until attaining a particular system performance. Before partitioning the data, it is recommended to shuffle the dataset to avoid the biasing of the results.

B.3. Hyperparameter optimization methods

Hyperparameter optimisation is an important bottleneck in machine learning. Since every scenario has its features, the adjustment of the machine learning algorithm requires an intensive test to find the configuration that yields the best performance. To overcome this problem, two alternatives are typically used: random search and grid search. On the one hand, random search consists of selecting a group of hyperparameters at random and testing the system's performance with this configuration. Once tested, the method randomly selects another group of hyperparameters and tests the system's performance again. This process loops the number of times chosen by the designer. Finally, the set of hyperparameters used is the one that yielded the best performance.

The random search method works fine when the number of hyperparameters is too large to combine them. However, the obtention of the best hyperparameters set is very dependent on the number of sets tested. Besides, the system never reaches the optimal solution. As an alternative, the grid search method proposes to combine all the hyperparameters among them to find the most optimal combination. Although it is a time demanding process that exponentially increases with the number of hyperparameters, the best hyperparameter solution is guaranteed.

UML diagrams

Figure C.1 shows the unified modelling language diagram for the decision-making architecture developed in this thesis.

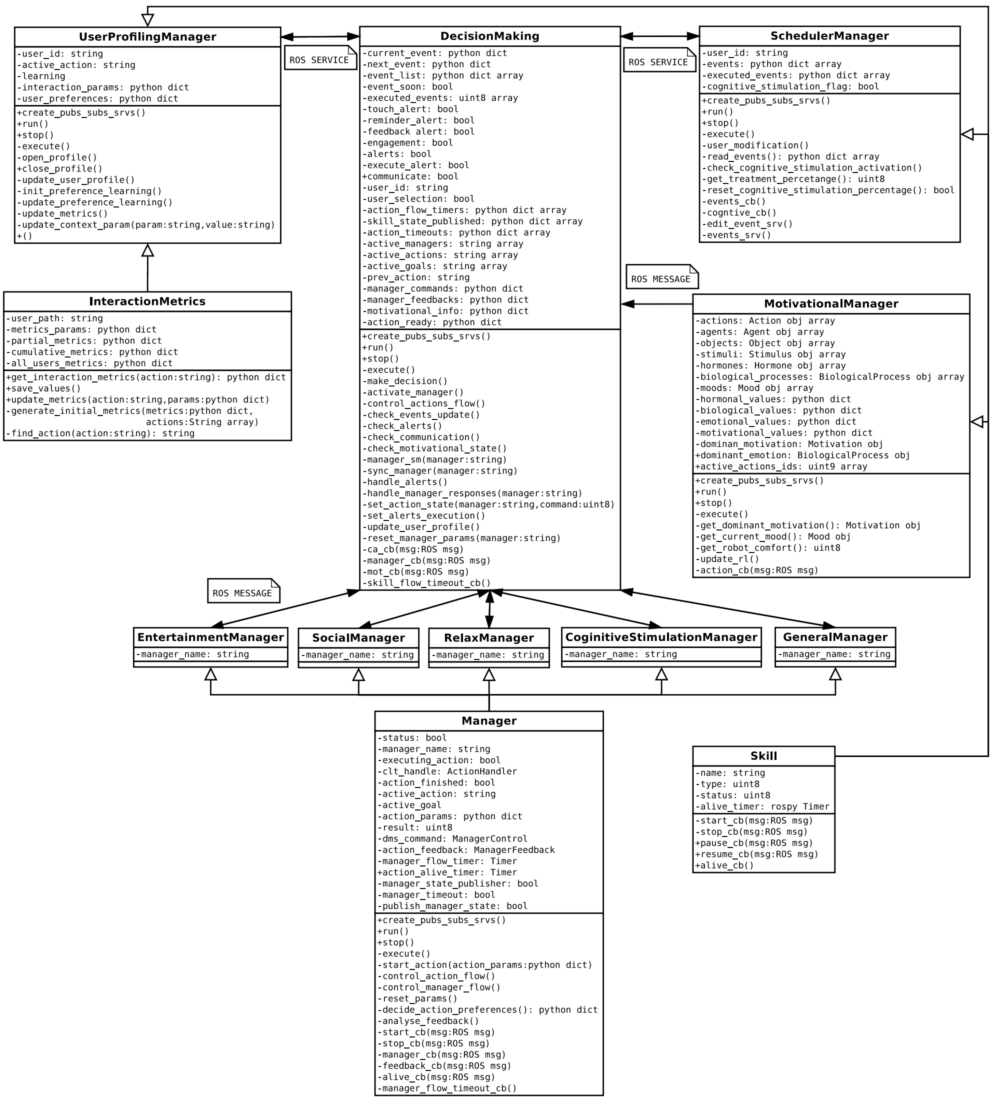


Figure C.1: Decision-making system UML diagram.

Figure C.2 shows the unified modelling language diagram regarding the motivational model developed in this thesis.

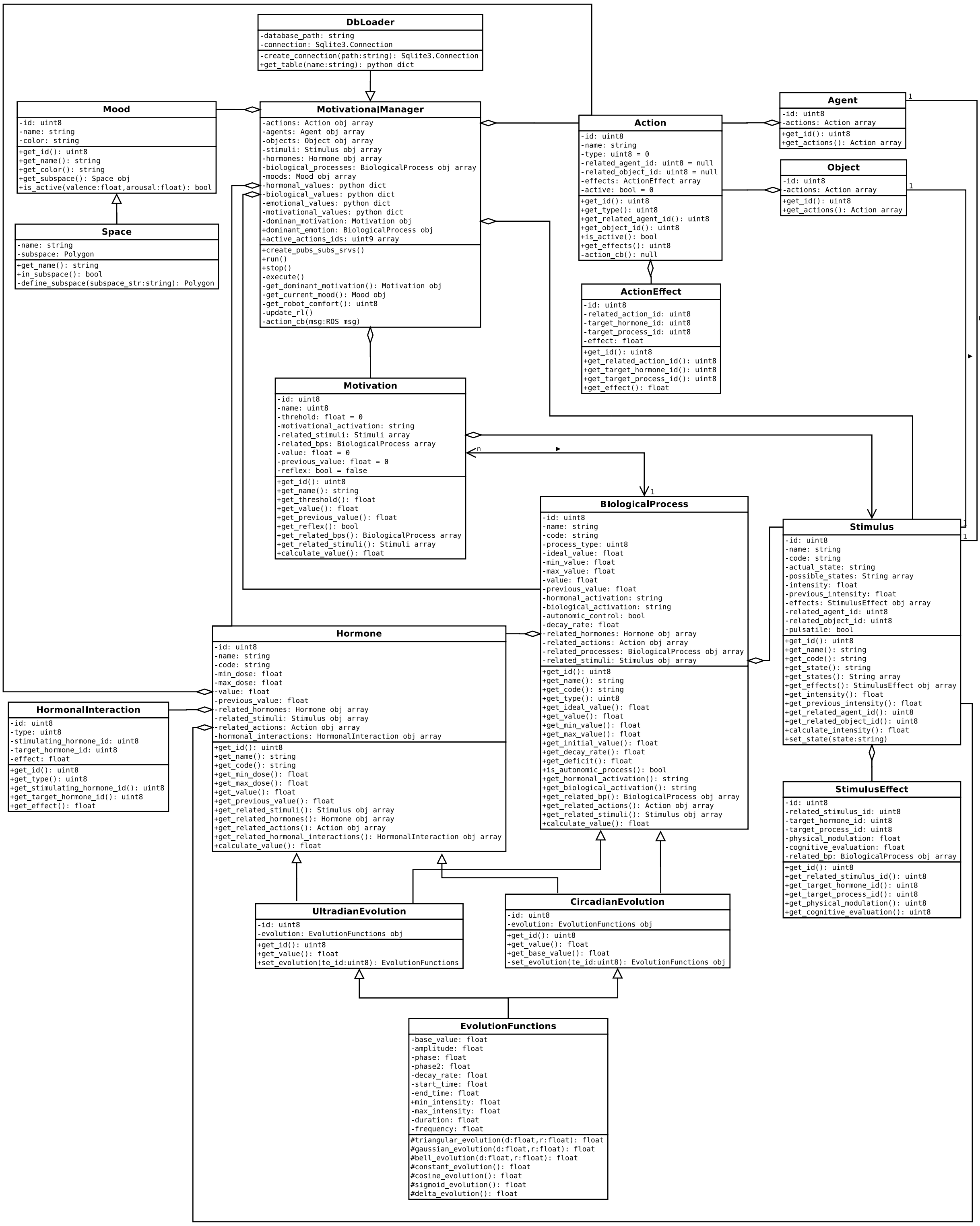


Figure C.2: Motivational model UML diagram.

APPENDIX D

Questionnaires

These appendices contain the questionnaires used to evaluate the HRI experiments conducted for testing the different functionalities developed in this thesis.

D.1. User profiling questions

Following, we list the questions used for retrieving user information through HRI.

Personal questions:

- What is your name?
- What is your first surname?
- What is your second surname?
- What day is your birthday?
- In which month is your birthday?
- What year were you born?
- Where do you live?
- What is your country of origin?
- What is your occupation?
- What is your educational level?

Interest questions:

Are you interested in ...?

- Videogames
- Reading
- Photography

- Films
- TV series
- Cooking
- Social Networks
- News
- Weather
- Internet
- Music
- Videos
- Purchasing

Preferences questions:

How much do you like to ...

- Play quiz games?
- Play classic games?
- Play brain training games?
- Listen to English pop?
- Listen to English rock?
- Listen to Spanish pop?
- Listen to Spanish rock?
- Listen to Latin music?
- Watch animal photos?
- Watch amazing landscapes photos?
- Watch historical monuments photos?
- Be informed about last news?
- Be informed about national news?
- Be informed about international news?
- Be informed about sports news?
- Be informed about the weather forecast?

- Watch comedy videos?
- Watch cooking recipes videos?
- Watch videos about epic sports moments?
- Listen to classical audiobooks?
- Listen to historical facts?
- Listen to tales?
- Listen to funny jokes?
- Listen to Spanish sayings?

D.2. Ad-hoc questions

In this section we state those ad-hoc questions used to evaluate some of the HRI experiments carried out in this thesis.

Engagement question:

- How much entertained are you perceiving this process? (¿Cúanto de entretenido te está pareciendo este proceso?)

Personalisation perceived:

- **Q1:** How appropriate did you find the first activity selected by the robot according to your preferences? (¿Cómo de apropiada encontraste la primera actividad según tus preferencias?)
- **Q2:** How appropriate did you find the second activity selected by the robot according to your preferences? (¿Cómo de apropiada encontraste la segunda actividad según tus preferencias?)
- **Q3:** How appropriate did you find the third activity selected by the robot according to your preferences? (¿Cómo de apropiada encontraste la tercera actividad según tus preferencias?)
- **Q4:** In general, do you think the activities proposed by the robot are adequate to you? (En general, ¿crees que las actividades propuestas por el robot se adecuan a ti?)
- **Q5:** Have you noticed that the robot knows information about yourself? (¿Has percibido que el robot sabe información sobre ti?)
- **Q6:** Have you noticed that the robot behaviour adapts to yourself? (¿Has percibido que el comportamiento del robot se adapte a ti?)

D.3. Robotic Social Attributes Scale (RoSAS)

The RoSAS [359] attributes scales measures the users' perception towards the categories *Warmth*, *Competence*, and *Discomfort*. Each category is composed by the six attributes shown below (with their Spanish translation):

Warmth (Cordialidad):

- Reliable (Fiable)
- Competent (Competente)
- Knowledgeable (Culto)
- Interactive (Interactivo)
- Responsive (Activo)
- Capable (Hábil)

Competence (Capacidad):

- Organic (Natural)
- Sociable (Sociable)
- Emotional (Emocional)
- Compassionate (Compasivo)
- Happy (Contento)
- Feeling (Sensible)

Discomfort (Incomodidad):

- Awkward (Incómodo)
- Scary (Siniestro)
- Strange (Extraño)
- Awful (Horrible)
- Dangerous (Peligroso)
- Aggressive (Agresivo)

D.4. System Usability Scale (SUS)

The SUS questionnaire [335] the usability of a system using ten questions that the participants had to answer using a 5-point Likert Scale. Then, the ratings of these questions are aggregated into a final score that represents the system usability from 0 to 100 units. Table D.1 shows the SUS questions and their Spanish translation.

Question	Spanish translation
I think that I would like to use this robot frequently	Me gustaría usar el robot más frecuentemente
I found the robot unnecessarily complex	Encuentro al robot demasiado complejo
I thought the robot was easy to use	Pienso que el robot es fácil de usar
I think that I would need the support of a technical person to be able to use this robot	Pensaba que necesitaría ayuda de otra persona para usar el robot
I found the various functions in this robot were well integrated	Las diferentes funciones del robot están bien integradas
I thought there was too much inconsistency in this robot	Creo que el robot resulta inconsistente
I would imagine that most people would learn to use this robot very quickly	Creo que la mayoría de la gente aprenderá a usar el robot rápidamente
I found the robot very cumbersome to use	El proceso de interacción me ha parecido pesado de completar.
I felt very confident using the robot	Me he sentido seguro completando este proceso.
I needed to learn a lot of things before I could get going with this robot	Necesité aprender muchas cosas nuevas antes de poder responder al robot

Table D.1: System Usability Scale questions and their Spanish translation.

D.5. Godspeed series

The Godspeed series [348] is a questionnaire that measures how robot users perceived the *Anthropomorphism*, *Animacy*, *Likeability*, *Intelligence*, and *Security*. Each attribute is formed by 3 to 6 items whose rating (obtained using a Likert scale) faces two opposite terms. Finally, the scores of the items are aggregated using an average. Next, we state the Godspeed attributes and items in English and Spanish.

Anthropomorphism (Antropomorfismo):

- Fake-Natural (Falso-Natural)
- Machinelike-Humanlike (Con aspecto de máquina-Con aspecto humano)
- Unconscious-Conscious (Inconsciente-Consciente)
- Artificial-Lifelike (Artificial-Vivo)
- Moving rigidly-Moving elegantly (Se mueve con rigidez-Se mueve con fluidez)

Animacy:

- Dead-Alive (Inerte-Con vida)
- Stagnant-Lively (Inactivo-Vivaz)
- Mechanical-Organic (Mecánico-Orgánico)

- Artificial-Lifelike (Artificial-Parece vivo)
- Inert-Interactive (Estático-Interactivo)
- Apathetic-Responsive (Indiferente-Atento)

Likeability:

- Dislike-Like (Digusta-Gusta)
- Unfriendly-Friendly (No amigable-Amigable)
- Unkind-Kind(Descortés-Cortés)
- Unpleasant-Pleasant (Desagradable-Agradable)
- Awful-Nice (Feo-Bonito)

Perceived intelligence:

- Incompetent-Competent (Incompetente-Competente)
- Ignorant-Knowledgeable (Ignorante-Culto)
- Irresponsible-Responsible (Irresponsable-Responsable)
- Unintelligent-Intelligent (Sin inteligencia-Inteligente)
- Foolish-Sensible (Insensato-Juicioso)

Perceived security:

- Anxious-Relaxed (Ansioso-Relajado)
- Calm-Agitated (Tranquilo-Agitado)
- Quiescent-Surprised (No sorprendido-Sorprendido)