Chapter 2
Modeling Emotions in Robotic Socially Believable Behaving Systems

Anna Esposito and Lakhmi C. Jain

Abstract This book aims to investigate the features that are at the core of human interactions to model the involved emotional processes, in order to design and develop autonomous systems and algorithms able to detect early signs of changes, in moods and emotional states. The attention is focused on emotional social features and the human's ability to decode and encode emotional social cues while interacting. In order to do this, the book will propose a series of investigations that gather behavioral data from speech, handwriting, facial, vocal and gestural expressions. This is done through the definition of behavioral tasks that may serve to produce changes in the perception of emotional social cues. Specific scenarios are designed to assess users' emphatic and social competencies. The collected data are used to gain knowledge on how behavioral and interactional features are affected by individuals' moods and emotional states. This information can be exploited to devise multidimensional models of multimodal interactional features that will serve for measuring the degree of empathic relationships developed between individuals and allow the design and development of cost-effective emotion-aware technologies to be used in applicative contexts such as remote health care services and robotic assistance.

A. Esposito (✉)
Dipartimento di Psicologia and IIASS, Seconda Università di Napoli, Caserta, Italy
e-mail: iiass.annaesp@tin.it

L.C. Jain
Bournemouth University, Bournemouth, UK
e-mail: jainlc2002@yahoo.co.uk

L.C. Jain
University of Canberra, Canberra, ACT 2601, Australia

2.1 Introduction

The realization of a robotic agent capable of natural interactions with humans raises a number of issues and problems related to (a) the acquisition of sufficient competences to afford a full description and understanding of the user's perspective, and thus the creation of reliable and functional user models, (b) the agent's appearance and design, including the to be implemented interface, and (c) the reliability and credibility of a such artificial agent under the variety of application areas its tasks are devoted to [7, 12, 14]. The application areas may vary from training/education systems, to human-computer interfaces, entertainment, computer games, and several more. Each may require different agent's abilities, at a different level of complexity, ranging from natural speech communication to autonomous behaviors, detection and interpretation of user moods, personalities, and emotional behaviors, adaptation to the physical, organizational, and socio-cultural context, as well as, user needs and requirements.

Robotic socially believable behaving agents must be adaptive systems programmed to interact with humans by simulating typical spontaneous human interactions. To this aim it is needed a comprehensive model of the possible sets of "user" final conducts (the "user model"), as well as, in order for the agent to properly act, it is necessary to provide a proprioceptive description of the agent itself (the "model of itself", [5]. In human-machine interaction, a user model is what the system "knows " about the perspective and/or the expectations of the actual user, normally including both psychological data and a user profile that estimates the preferences of a specific class or a wide range of users. Psychological data tend to function effectively as a paradigm for a wider set of users, and are a reliable source for modeling the user's prospective [24]. For example, in the case of agents devoted to act a "butler" for guiding users in a foreign city, the user model needs to provide a comprehensive inventory of possible human requests and actions/reactions to expected and unexpected situations in terms of the reliability, confidence, facility of interaction/communication, as well as, trustworthiness, satisfaction, and credibility the users ascribe to the agent under different circumstances. Moreover, together a model of the possible user behavioral response, it is also important to anticipate such responses by means of a motivated design in order to minimize negative effects that can produce the user rejection of the agent.

Since emotions play a very important role in many aspects of our lives, including decision making, perception, learning, and behavior, and emotional skills are an important component of intelligence, especially for human-human interactions, the next generation of cognitive architectures must integrate and incorporate principles of emotionally colored interactions to define realistic models of human-machine interaction and suggest novel computational approaches for the implementation of real-time believable, autonomous, adaptive, and context-aware robotic agents [10, 13]. To this aim, fair efforts have been devoted to recognize and synthesize human emotional states exploiting different modalities, including: speech, facial expressions, gestures, and physiological signals as EEG, ECG, and others [1–3, 20–23, 26]. Some works investigated on the possibility to combine signals from multiple

sources, with the general aim of improving the emotion classification accuracy and synthesis [6, 11]. However, a truly cognitive real-time multimodal system capable of working with spontaneous emotional material is still missing. The chapters proposed in this book aim to further progress in this direction, even though automatic synthesis and recognition of human emotions remain an example of computational task that cannot be perfectly solved by classical pattern recognition and machine learning algorithms.

2.2   Content of the Book

Sophisticated and functional computational instruments able to recognize, process, store, and synthesize relevant emotional signals, as well as interact with people, displaying reactions that show abilities of appropriately sensing and understanding (under conditions of limited time) environmental changes and producing suitable, autonomous, and adaptable responses created great promises in Information Communication Technology (ICT). However, Human Computer Interaction (HCI) paradigms that account for users' emotional states and emotional reactions are still far from being implemented. This book is signaling advances in these directions taking into account the multifunctional role of emotions in driving communication and interactional exchanges. To this aim, the book includes nine investigations on the role of emotions in shaping intentions, motivations, learning, and decision making. The first chapter by Vernon et al. [25] affords the problem of modeling (in socially believable behaving systems) actions, attentions, goals, and intentions, particularly intentions considered by the authors as the ability of the agent to read others minds and understand others' perspectives and beliefs. In this context, emotions will drive the efficient behavior of the agents, since as stated by the authors, the perception of social stimuli produces "*bodily states in the perceiving agent* [and] *trigger* [in it] *affective states* [which in turn affect] *the agent's physical and cognitive performance*" [25, p. 3]. Corrigan et al. [8] faces similar problems from an implementation point of view considering aspects of human engagements with robotic and virtual agents, which are still actions guided by intentions and abilities to understand others' perspectives. The contribution of Belpaeme et al. [4] approaches similar problematics at a developmental level. The authors show that "*cognition emerges from the interaction between the brain, the body and the physical environment*" exploiting the iCub humanoid platform. The interesting paradigm emerging from their results is that "*artificial cognition, just as its natural counterpart, benefits from being grounded and embodied*" in a body, a brain, a physical, social context. The missing of one of these four constituents still allow to create artificial cognition, however, the complete system seems to be more efficient and effective. Meudt's et al. [19] contribution intend to show that the ability of an agent to recognize users' emotions will facilitate the users' adaptation process and improve the human interaction with the agent, making its function as companion or assistive technology more reliable. The contribution of Lewandowska-Tomaszczyk and Wilson [17] underline the physical and

moral role of the disgust across different cultures, showing how important is the social setting and why "*an emotion-sensitive socially interacting robot would need to encode and decode* [such emotion] *in order to competently* and appropriately interact with the environmental culture [17, p. 1]. On a similar theme is the contribution of Maricchiolo et al. [18], which affords the analysis of nonverbal (gestural) and physiological (heart rate and skin conductance) reactions to disagreeable (disgusting?) messages. Surely very far from disgust is the contribution of Dupont et al. [9] which reports on the result of a four-year EU project investigation on the "laughter": the ILHAIRE Project (http://www.ilhaire.eu/project). The authors describe the collected data and the multi-determined role of the "laughter" in social interaction. Finally, the last two contributions of this book by Hunyadi et al. [16] and Gangamohan et al. [15] are dedicated to emotional speech and in particular to emotional prosodic features extracted from the HuComTech Corpus and all the emotional speech features able to recognize emotional vocal expressions.

2.3 Conclusions

The readers of this book will get a taste of the major research areas in modeling emotions and of the multifunctional role of emotions in generating actions, goals, attentions, and intentions, as well as on the different paradigms to model emotions by analyzing interactional exchanges. The research topics afforded by the book cover research fields related to psychology, sociology, philosophy, computer science, robotics, signal processing and human-computer interaction. The contributors to this volume are leading authorities in their respective fields. The book captures and presents interesting aspects of communicative exchanges and is fundamental in studies, such as robotics, where multidisciplinary facets need to be considered in order to succeed in the implementation of robotic companions, and assistive technologies. In particular the book covers aspects of emotional information processing during interactional exchanges which would lead to the implementation of socially believable, autonomous, adaptive, context-aware situated HCI systems.

References

1. Atassi H, Esposito A (2008) Speaker independent approach to the classification of emotional vocal expressions. In: Proceedings of IEEE conference on tools with artificial intelligence (ICTAI 2008), vol 1. Dayton, 3–5 Nov 2008, pp 487–494
2. Atassi H, Esposito A, Smekal Z (2011) Analysis of high-level features for vocal emotion recognition. In: Proceedings of 34th IEEE international conference on telecom and signal processing (TSP), Budapest, 18–20 Aug 2011, pp 361–366
3. Atassi H, Riviello MT, Smékal Z, Hussain A, Esposito A (2010) Emotional vocal expressions recognition using the cost 2102 italian database of emotional speech. In: Esposito A et al. (eds)

Development of multimodal interfaces: active listening and synchrony, LNCS 5967, Springer, Berlin, pp 255–267

4. Belpaeme T, Adams S, De Greeff J, Di Nuovo A, Morse A, Cangelosi A (2016) Social development of artificial cognition. This volume

5. Benyon D, Turner P, Turner S (2005) Designing interactive systems: people, activities, contexts, technologies. Pearson Education, Harlow

6. Castellano G, Kessous L, Caridakis G (2008) Emotion recognition through multiple modalities: face, body, gesture, speech. Affect and emotion in human-computer interaction. Springer, Berlin, pp 92–103

7. Cordasco G, Esposito M, Masucci F, Riviello MT, Esposito A, Chollet G, Schlögl S, Milhorat P, Pelosi G (2014) Assessing voice user interfaces: the vAssist system prototype. In: Proceedings of the 5th IEEE international conference on cognitive infocommunications, Vietri sul Mare, 5–7 Nov 2014, pp 91–96

8. Corrigan LJ, Peters C, Küster D, Castellano G (2016) Engagement perception and generation for social robots and virtual agents. This volume

9. Dupont S, Çakmak H, Curran W, Dutoit T, Hofmann J, McKeown G, Pietquin O, Platt T, Ruch W, Urbain J (2016) Laughter research: a review of the ILHAIRE project. This volume

10. Esposito A (2013) The situated multimodal facets of human communication. In Rojc M, Campbell N (Eds), Coverbal synchrony in human-machine interaction, chap. 7. CRC Press, Taylor & Francis Group, Boca Raton, pp 173–202

11. Esposito A, Esposito AM (2012) On the recognition of emotional vocal expressions: motivations for an holistic approach. Cogn Process 13(2):541–550

12. Esposito A, Fortunati L, Lugano G (2014) Modeling emotion, behaviour and context in socially believable robots and ICT interfaces. Cogn Comput 6(4):623–627

13. Esposito A, Esposito AM, Vogel C (2015) Needs and challenges in human computer interaction for processing social emotional information. Patter Recognit Lett 66:41–51

14. Fortunati L, Esposito A, Lugano G (2015) Beyond industrial robotics: social robots entering public and domestic spheres. Inf Soc: Int J 31(3):229–236

15. Gangamohan P, Kadiri SR, Yegnanarayana B (2016) Analysis of emotional speech: A review. This volume

16. Hunyadi L, István Szekrényes I, Kiss H (2016) Prosody enhances cognitive infocommunication: materials from the HuComTech corpus. This volume

17. Lewandowska-Tomaszczyk B, Wilson PA (2016) Physical and moral disgust with socially believable behaving systems in different cultures. This volume

18. Maricchiolo F, Gnisci A, Cerasuolo M, Ficca, Bonaiuto M (2016) Speaker's hand gestures can modulate receiver's negative reactions to a disagreeable verbal message. This volume

19. Meudt S, Schmidt-Wack M, Honold F, Schüssel F, Michael Weber M, Schwenker F, Palm G (2016) Going further in affective computing: how emotion recognition can improve adaptive user interaction. This volume

20. Milhorat P, Schlögl S, Chollet G, Boudy J, Esposito A, Pelosi G (2014) Building the next generation of personal digital assistants. In: Proceedings of the 1st IEEE international conference on advanced technologies for signal and image processing - ATSIP'2014, Sousse, 17–19 March 2014, pp 458–463

21. Placidi G, Avola D, Petracca A, Sgallari F, Spezialetti M (2015) Basis for the implementation of an EEG-based single-trial binary brain computer interface through the disgust produced by remembering unpleasant odors. Neurocomputing 160:308–318

22. Ringeval F, Eyben F, Kroupi E, Yuce A, Thiran JP, Ebrahimi T, Lalanne D, Schuller B (2014) Prediction of asynchronous dimensional emotion ratings from audiovisual and physiological data. Pattern Recogn Lett Elsevier 66(C):22–30

23. Schuller B (2015) Deep learning our everyday emotions: a short overview. In: Bassis S et al (eds) Advances in neural networks: computational and theoretical issues, vol 37. SIST Series, Springer, Berlin, pp 339–346

24. van der Veer GC, Tauber MJ, Waern Y, van Muylwijk B (1985) On the interaction between system and user characteristics. Behav Inf Technol 4:284–308

25. Vernon D, Thill S, and Ziemke T (2016) The role of intention in cognitive robotics. This volume
26. Vinciarelli A, Esposito A, André E, Bonin F, Chetouani M, Cohn JF, Cristan M, Fuhrmann F, Gilmartin E, Hammal Z, Heylen D, Kaiser R, Koutsombogera M, Potamianos A, Renals S, Riccardi G, Salah AA (2015) Open challenges in modelling, analysis and synthesis of human behaviour in human-human and human-machine interactions. Cogn Comput 7(4):397–413