Migratable AI Systems for Tailoring User Experience through Multimodal Affective-Cognitive State Analysis

Ludovica La Monica[1,2,*]

[1]Institute of Cognitive Science and Technology, National Research Council, 00185 Rome, Italy
[2]Unit of Computer Systems and Bioinformatics, Department of Engineering, Università Campus Bio-Medico di Roma, 00128 Rome, Italy

Abstract

In the era of digital technologies permeating numerous aspects of society, the development of intelligent artificial agents endowed with autonomy, social capabilities, reactivity, and proactivity has become a pivotal innovation. Traditionally, these agents were confined to singular forms, known as *embodiments*, which dictated their capabilities and the environments in which they operated. To transcend these limitations and enable agents to expand their capabilities, the concept of *agent migration* has emerged, allowing them to seamlessly transition between different embodiments. My research program delves into the realm of agent migration, uncovering its challenges and potential opportunities. The central objective is to leverage this technology to deepen our understanding of human affective and cognitive states across various contexts. Specifically, my aim is to develop a migratory agent architecture capable of recognizing and characterizing individuals' emotional experiences using different sets of capabilities tied to distinct embodiments. In summary, the intent of this research is to harness migratable AI to promote the technical advantages brought about by agent migration and gain insights into how to utilize the acquisition of additional capabilities resulting from migration to craft a more comprehensive and potentially enduring architecture, to be applied in the realm of affective computing.

Keywords

Intelligent Agent, Embodied Agent, Embodied Cognition, Migratable AI, Affective Computing, Emotion recognition

1. Introduction

Nowadays, an increasing number of human activities rely on digital technologies in the current society [1]. This phenomenon has driven innovation in the field of artificial agents, i.e., sophisticated software entities exhibiting specific characteristics such as autonomy, social abilities, reactivity, and proactivity [2]. Traditionally, such agents were constrained to a single form, known as *embodiment*, representing their physical or digital manifestation. Embodiment plays a key role in the development of an intelligent agent as it determines the *sense of presence* that allows users to recognize it and establish a relationship with it [3]. Simultaneously, embodiment dictates the intrinsic capabilities of the agent and the environment in which it can operate. This

AIxIA'23: 22nd International Conference of the Italian Association for Artificial Intelligence, November 06–09, 2023, Rome, Italy

*Corresponding author.
ludovica.lamonica@unicampus.it (L. La Monica)

concept implies that the agent's capabilities are inherently tied to and constrained by its form. To overcome this limitation and enable agents to extend their abilities, the concept of *agent migration* has been introduced. This approach allows agents to transfer their *"being"* across different embodiments, enabling them to acquire the capabilities associated with each of them and possibly expand their interactions with the surrounding environment and users. Several studies explored aspects related to agent migration, investigating the advantages and challenges associated with this practice. Agent migration, by and large, is defined as the process where intelligent agent moves through different embodiments, remaining active in one at a time [4].

The concept of an agent capable of migration allows for the separation of the agent's mind from the body it occupies [5]. The mind defines its essence, being an intelligent system that makes decisions and plans actions to be executed by the body. The body is seen as a shell that carries out the actions prescribed by the mind and determines the capabilities it can leverage. The detachment of the mind from a specific body not only facilitates the acquisition of extended capabilities but eliminates the limitation of environmental confinement, facilitating uninterrupted interaction with the agent, as the mind can accompany the user across diverse contexts by merely changing its physical form.

The objective of my Ph.D. research program (started this year) is to examine the concept of agent migration, exploring its challenges and opportunities, with the aim of harnessing this technology to enhance the comprehension of the affective and cognitive states of individuals across various contexts. Specifically, the goal is to develop a migratory agent architecture that enables the recognition and comprehensive characterization of subject's emotional experiences, utilizing different sets of capabilities based on the embodiment it assumes.

2. Review of Literature

Focus on Migratable AI. Migrant agents and the migration process have been studied in the context of social interactions and human-machine interactions, i.e., HRI (Human-Robot Interaction) and HCI (Human-Computer Interaction). Pioneering this field, Imai et al. (1999) addressed agent migration with the goal of enhancing human-robot interaction [6]. In their work, they introduced the migratable system ITAKO, which allowed an intelligent agent to move between a mobile PC and an autonomous robot.

In the 2000s, the UCD (University College Dublin) team introduced the Agent Chameleons architecture with the aim of enabling seamless transitions between physical and virtual information environments, emphasizing the crucial role of migration as a means for expanding the capabilities of autonomous systems [3].

Subsequently, research began to consider the effects of migration on users. Indeed, while migration enhances agent capabilities and frees them from environmental constraints, the critical issue of user recognition of the agent's identity emerged. Studies, such as that of Gomes et al. [4], analyzed how users perceive the migration process and the number of perceived agent identities. Concurrently, Holz et al. [1] evaluated different forms of embodiment (i.e., physical or virtual) to optimize agent capabilities, particularly in relational terms. Koay et al. [7] looked at the use of visual representation of migration as a method of retaining the agent's identity and improving its identification, so facilitating the user's interaction. In alignment with these

approaches, Tejwani et al. [8] proposed a study aimed at assessing the effect on user-agent interaction quality resulting from maintaining specific parameters during the migration process, namely agent identity and acquired information.

It is evident, therefore, that the agent's identity constitutes a key feature in the migration process, fostering the establishment of stable and enduring relationships between the user and the agent.

Focus on Emotion Recognition. Numerous attempts have been made to classify emotions in order to understand human behavior [9]. The number of categories in which emotions should be classified has always been a contentious topic. The two main recognized theories are Ekman's discrete theory of emotions [10] and Russell's dimensional theory of emotions [11]. Systems for recognizing emotional states leverage these theories to characterize acquired information and establish connections between emotions, affect, and behavior.

One of the channels used to gather emotional information is speech. In this context various approaches are available depending on the audio signal format. These data can be processed as raw audio waveforms or in 2D format, such as spectrograms [12]. When analyzing raw audio waveform data, a widely used approach involves the use of the WaveNet architecture [13, 14]. On the other hand, deep learning technologies, such as Long Short-Term Memory (LSTM) artificial neural networks integrated with Convolutional Neural Networks (CNN), have been employed for the analysis of 2D audio signals [15, 16, 17]. Another relevant methodology involves the use of transformers [18].

Facial expressions, similar to natural language, are a manifestation of human emotions. The Facial Action Coding System (FACS) [19], developed by Ekman and Friesen in 1978, is the most well-known and widely used system for facial activity analysis. More recently, with the introduction of Graphics Processing Units (GPUs) and Convolutional Neural Networks, real-time applications have been developed, such as Microsoft service Emotion API [20] and the MIT Media Lab Affectiva technology Affedex [21].

From a physiological perspective, emotions leave traces that can be studied. Signals such as electrodermal activity, also known as Galvanic Skin Response (GSR), heart and respiratory rate, brain activity, and many others are typically considered when attempting emotion classification through a dimensional approach [22].

In conclusion, it is evident that emotion detection and, more broadly, Affective Computing are fields of great interest, despite the ongoing challenges.

3. Problem Statement

The literature review illustrates that agent migration is not devoid of challenges and complications. One of the primary challenges concerns the continuous recognition of the agent by the user, as the various embodiments may not be immediately attributed to a singular entity. This has prompted fundamental inquiries, which have undergone comprehensive scrutiny, regarding the nature of the agent's identity and the mechanisms facilitating recognition when its form may undergo changes. Indeed, the quest for optimal interaction between humans and agents has constituted the central focus of recent research in the domain of migratable AI. Simultaneously,

it becomes imperative to understand how to leverage migratable AI, what technical benefits agent migration can bring and how to use the acquisition of additional capabilities brought about by migration to produce a more comprehensive and potentially enduring architecture. These fundamental questions have served as the driving impetus behind my research endeavors in this field.

Another key aspect in the designing of the research project was the recognition of the significance of analyzing individuals' emotional states. The conducted assessment revealed that emotions represent a fundamental form of non-verbal interaction, and the ability to express emotional states prompted by everyday situations is a crucial aspect of human life. Simultaneously, the capacity to recognize emotional states expressed by others plays a pivotal role in both interpersonal relationships and the psychological well-being of individuals. Affective computing constitutes the main research area dedicated to the analysis of these emotional states: it leverages information acquired through the analysis of specific signals, such as facial expressions, natural language, body gestures, and physiological cues, to infer individuals' affective-cognitive states. However, research has underscored that information such as natural language and facial expressions, when considered in isolation, may lead to erroneous interpretations, as they can be consciously controlled by humans who may choose to conceal their emotional states. Conversely, physiological signals represent a more objective and reliable type of information, albeit more challenging to capture and, most importantly, interpret. These considerations have prompted the idea of applying migratable artificial intelligence to the field of affective computing.

4. My work

The main objective of this research is to develop an intelligent agent capable of migrating between different embodiments in order to leverage a broader range of capabilities and assist the user across various types of tasks. The capabilities of this intelligent system to migrate will enable the consistent and accurate integration of various type of information gathered during user interaction.

The proposed project aims to harness these capabilities to enhance the recognition of emotions expressed by individuals. Specifically, the agent's ability to migrate across different embodiments will enable it to acquire, analyze, and combine various types of user-related information. Consequently, the recognized emotional state will no longer rely on a single information channel but will result from a comprehensive analysis involving diverse data sources.

5. Conclusion

This work delves into the concept of agent migration, shedding light on both its advantages and the remaining challenges. Specifically, it identifies migration as the means to enhance the versatility and adaptability of intelligent agents across various contexts.

The main objective of this research is to expand the range of capabilities within the intelligent system, enabling it to assist users in diverse tasks. The focal point of this endeavor lies in the analysis of emotions and affective computing, recognizing the paramount importance of emotional state analysis in human interactions. Consequently, the aim is to utilize migratory

capabilities to integrate diverse user-related information, ultimately enhancing the capacity to recognize individuals' emotional states.

In conclusion, this research seeks to harness migratory artificial intelligence, identify the technical advantages stemming from agent migration, and acquire insights on how to leverage the acquisition of additional capabilities brought about by migration to develop more comprehensive and potentially enduring architectures in the field of affective computing.

Acknowledgments

This research is carried out as PhD student enrolled in the National PhD in Artificial Intelligence for Health and Life Sciences, XXXVIII Bis cycle, Università Campus Bio-Medico di Roma.

References

[1] T. Holz, M. Dragone, G. M. O'Hare, Where robots and virtual agents meet: a survey of social interaction research across milgram's reality-virtuality continuum, International Journal of Social Robotics 1 (2009) 83–93.

[2] M. Wooldridge, N. R. Jennings, Intelligent agents: Theory and practice, The knowledge engineering review 10 (1995) 115–152.

[3] G. M. O'Hare, B. R. Duffy, Agent chameleons: Migration and mutation within and between real and virtual spaces, The Society for the Study of Artificial Intelligence and the Simulation of Behaviour AISB (2002) 3–5.

[4] P. F. Gomes, A. Sardinha, E. Márquez Segura, H. Cramer, A. Paiva, Migration between two embodiments of an artificial pet, International Journal of Humanoid Robotics 11 (2014) 1450001.

[5] K. Hassani, W.-S. Lee, On designing migrating agents: from autonomous virtual agents to intelligent robotic systems, in: SIGGRAPH Asia 2014 Autonomous Virtual Humans and Social Robot for Telepresence, 2014, pp. 1–10.

[6] M. Imai, T. Ono, T. Etani, Agent migration: communications between a human and robot, in: IEEE SMC'99 Conference Proceedings. 1999 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No. 99CH37028), volume 4, IEEE, 1999, pp. 1044–1048.

[7] K. L. Koay, D. S. Syrdal, M. L. Walters, K. Dautenhahn, A user study on visualization of agent migration between two companion robots (2009).

[8] R. Tejwani, F. Moreno, S. Jeong, H. W. Park, C. Breazeal, Migratable ai: Effect of identity and information migration on users' perception of conversational ai agents, in: 2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), IEEE, 2020.

[9] A. Dzedzickis, A. Kaklauskas, V. Bucinskas, Human emotion recognition: Review of sensors and methods, Sensors 20 (2020) 592.

[10] P. Ekman, et al., Basic emotions, Handbook of cognition and emotion 98 (1999) 16.

[11] J. A. Russell, A circumplex model of affect., Journal of personality and social psychology 39 (1980) 1161.

[12] M. Zielonka, A. Piastowski, A. Czyżewski, P. Nadachowski, M. Operlejn, K. Kaczor, Recognition of emotions in speech using convolutional neural networks on different datasets, Electronics 11 (2022).
[13] A. v. d. Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbrenner, A. Senior, K. Kavukcuoglu, Wavenet: A generative model for raw audio, arXiv preprint arXiv:1609.03499 (2016).
[14] S. K. Pandey, H. S. Shekhawat, S. Prasanna, Emotion recognition from raw speech using wavenet, in: TENCON 2019-2019 IEEE Region 10 Conference (TENCON), IEEE, 2019, pp. 1292–1297.
[15] J. Wang, M. Xue, R. Culhane, E. Diao, J. Ding, V. Tarokh, Speech emotion recognition with dual-sequence lstm architecture, in: ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2020, pp. 6474–6478.
[16] C.-W. Huang, S. S. Narayanan, Deep convolutional recurrent neural network with attention mechanism for robust speech emotion recognition, in: 2017 IEEE international conference on multimedia and expo (ICME), IEEE, 2017.
[17] G. Trigeorgis, F. Ringeval, R. Brueckner, E. Marchi, M. A. Nicolaou, B. Schuller, S. Zafeiriou, Adieu features? end-to-end speech emotion recognition using a deep convolutional recurrent network, in: 2016 IEEE international conference on acoustics, speech and signal processing (ICASSP), IEEE, 2016, pp. 5200–5204.
[18] G. Tamulevičius, G. Korvel, A. B. Yayak, P. Treigys, J. Bernatavičienė, B. Kostek, A study of cross-linguistic speech emotion recognition based on 2d feature spaces, Electronics 9 (2020) 1725.
[19] P. Ekman, W. V. Friesen, Facial action coding system, Environmental Psychology & Nonverbal Behavior (1978).
[20] A. Verma, D. Malla, A. K. Choudhary, V. Arora, A detailed study of azure platform & its cognitive services, in: 2019 International conference on machine learning, big data, cloud and parallel computing (COMITCon), IEEE, 2019, pp. 129–134.
[21] D. McDuff, A. Mahmoud, M. Mavadati, M. Amr, J. Turcot, R. e. Kaliouby, Affdex sdk: a cross-platform real-time multi-face expression recognition toolkit, in: Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems, 2016, pp. 3723–3726.
[22] L. Shu, J. Xie, M. Yang, Z. Li, Z. Li, D. Liao, X. Xu, X. Yang, A review of emotion recognition using physiological signals, Sensors 18 (2018) 2074.