6

Embodied Conversation:
A Personalized Conversational HCI Interface for Ambient Intelligence

Andrej Zgank

 https://orcid.org/0000-0002-1411-5293
University of Maribor, Slovenia

Izidor Mlakar
University of Maribor, Slovenia

Uros Berglez
University of Maribor, Slovenia

Danilo Zimsek
University of Maribor, Slovenia

Matej Borko
A1 Slovenija, Slovenia

Zdravko Kacic
University of Maribor, Slovenia

Matej Rojc
University of Maribor, Slovenia

INTRODUCTION

Ambient Intelligence (AmI) and Smart Environments (SmE) are built on three foundations: Ubiquitous computing, ubiquitous communication, and personalized adaptive interfaces. One of the key requirements is the fact that interaction takes place through natural interfaces, in such a way that people can perceive the presence of smart objects and machines as little as possible, and only when the system cannot achieve its functionality without intervening for the user's attention. Namely, the use of physical, especially non-freeform, controllers (e.g. touch-screen, remote, mouse, keyboard) hinders natural interface and establishes a strong barrier between the user and computer. Moreover, speech and body language are part of the input/output Human-Computer Interface (HCI) modalities, which have the benefit of being a natural modality for communication between humans. The speech as input modality can be used in many scenarios, which are also reflected in the complexity of the automatic speech recognition system. Humans are social beings, who utilize all available signals, from linguistics, language, and speech, to paralinguistic and non-verbal in the perception, understanding, and production of information in face-to-face communication. Some of the paralinguistic and non-verbal audio events can be processed in parallel with the automatic speech recognition system, with the possibility to share some modules, such as feature extraction.

Aggregation in (multimodal) understanding and use of natural language is especially supported through concepts of Conversational Interfaces (CIs) and Embodied Conversational Agents (ECA). In recent years, the CIs have become an important research and development topic in academia and the ICT industry. Smart environments utilize Artificial Intelligence (AI) increasingly and incorporate smart objects, such as: Devices, wearables, virtual agents, social robots, etc. Leading companies (e.g. Google, Amazon, Microsoft, Apple) have been making huge investments in the development of supporting technologies of Artificial Intelligence, such as Deep Learning and Natural Language Processing (NLP). The aim of these ventures is to create systems that will enable users of smart systems to obtain information and access services in a more natural, embodied conversational way. The classical vision of CIs incorporates dialog systems, and voice-enabled interfaces supported by Natural Language Processing. The multimodal vision of CIs, however, also aggregates embodiment and embodied cognition (as part of embodied language processing) in input (gesture-enabled user interfaces) and output (embodied conversational agents). Namely, CIs with virtual ECA capable of understanding verbal and non-verbal signals and capable of generating them, can engage with a human user on a more personal level. Particularly, through more conversation-like flow of the interaction, these Embodied Conversational Agents (supported by Natural Language Processing (NLP) and Embodied Language Processing (ELP)), may deliver new interaction models capable of adapting to the user's context and facilitating context of the conversational situation not only via modalities like speech, but also through visual interpretation and representation of information and social cues.

Despite all investment in this research area, an important amount of challenges still need to be solved. One barrier, which is connected with the users' desire to communicate in their native language, has to deal with the language questions. The languages outside the economic interest of leading companies are usually under-resourced, which presents an additional challenge for the intelligent ambient development process. A solution applied in the last few years is oriented toward collecting open and freely available language resources. In the case of such resources, some modalities, such as text, are available more frequently than others, as is speech combined with transcriptions. The availability of language resources is, thus, one of the key questions in the era of industry digitalization.

The intelligent ambient, and especially its manifestation in the form of the Embodied Conversational Agent, plays a crucial role in many areas of modern society. The application areas can range from general use in a smart home, to a dedicated use in a highly specialized e-Health environment. The common goal is to ease the user's interaction, with the objective to make it as natural as possible.

The background and literature review are presented first. The novel Ambient Intelligence embodied conversational model for HCI is proposed and the main modules, which are necessary to establish a natural communication, are described. The objective is to improve the quality of HCI in Ambient Intelligence, with the usage of advanced approaches. In addition, the services based on the proposed interaction model are presented. The chapter concludes with future research directions.

BACKGROUND AND LITERATURE REVIEW

A general overview of speech as the input modality is given in (Rabiner et al., 1993), where the authors addressed all different entities of an automatic speech recognition system. For more than two decades, Hidden Markov Models were used as basic acoustic models (Huang et al., 2010; Bilmes 2006) for different speech recognition scenarios that were applied in the research and industrial domains. Important progress was made in the area of robust speech processing, where noise reduction techniques and speech

Related Content

Considerations of Self in Recognising Prior Learning and Credentialing
Lloyd Hawkeye Robertsonand Dianne Conrad (2016). *Open Learning and Formal Credentialing in Higher Education: Curriculum Models and Institutional Policies (pp. 187-204).*
www.irma-international.org/chapter/considerations-of-self-in-recognising-prior-learning-and-credentialing/135646

Artificial Intelligence for Innovative Customer Services in the Hospitality Industry of India
Lakhvinder Singhand Dinesh Dhankhar (2024). *Leveraging AI and Emotional Intelligence in Contemporary Business Organizations (pp. 219-239).*
www.irma-international.org/chapter/artificial-intelligence-for-innovative-customer-services-in-the-hospitality-industry-of-india/335421

Strategies to Mentor Female Faculty: A Global Issue
Cassandra Sligh Conway, Yvonne Sims, Audrey McCrary-Quarles, Cynthia Salley Nicholson, Glacia Ethridge, Michelle Maultsby, Tammara Petrill Thomasand Susan Smith (2018). *Faculty Mentorship at Historically Black Colleges and Universities (pp. 126-150).*
www.irma-international.org/chapter/strategies-to-mentor-female-faculty/198829

Continuous Change in Educational Organizations
Yasar Kondakci, Merve Zayimand Kadir Beycioglu (2015). *Multidimensional Perspectives on Principal Leadership Effectiveness (pp. 305-322).*
www.irma-international.org/chapter/continuous-change-in-educational-organizations/121145

Shock Leadership: Leading Amidst Pandemics and Other Chaotic Change
Anton Shufutinsky, Bena Long, James R. Sibeland Darrell Norman Burrell (2021). *Global Perspectives on Change Management and Leadership in the Post-COVID-19 Era (pp. 136-159).*
www.irma-international.org/chapter/shock-leadership/274201