CHAPTER X

Human-Robot Interaction in the MOnarCH project

Víctor González-Pacheco and Álvaro Castro-González and María Malfaz1 and Miguel A. Salichs

RoboticsLab, Universidad Carlos III de Madrid vgonzale@ing.uc3m.es

This paper presents the design of the human-robot interaction (HRI) architecture developed in the context of the European project MOnarCH. In this project a team of robots operates in the pediatric ward of an oncological hospital. In this scenario the robots operate for long periods interacting with the different people they encounter: children, staff, etc. We present an HRI architecture that is capable of managing the different situations that the robots might face while interacting with people. We model these interactions as multimodal dialogues in which the robot has to perceive the current situation and express itself in a cohesive and coherent manner.

1 Introduction

In Human-Robot Interaction (HRI) we try to emulate with robots the way humans communicate and interact among them. Therefore, we need to provide a system that is able to deal with the conditions of the environment and the diversity of human responses. This system is in charge of customizing the way the robot communicates. In this paper we present the approach followed in the MOnarCH Project to handle the HRI. The aim of this project is to develop a networked system of social robots targeted to interact and to engage children who are hospitalized in the pediatrics ward of an oncological hospital in Lisbon, Portugal. The interaction activities of the robots will consist in, mainly, edutainment (education plus entertainment) activities. We propose an HRI architecture to handle the interactions during the operation of the robots in the hospital. This architecture is based on Communicative Acts (CAs), which are the minimum unit of communication between 2 entities (in this case, a robot and a human). The CA idea in inspired in the ideas of Searle (Searle, 1969) and others (Austin, 1975; Jakobson, 1960), but adapted to the MOnarCH scenarios.

To understand our approach, we first describe the MOnarCH hospital environment together with the robots that operate in such hospital (both in Section 2). Following, we describe in Section 3 the HRI architecture that enables the MOnarCH robots to adapt to the variety of social situations that occur in the hospital. We finally conclude the paper in Section 4 stating the major contributions of the MOnarCH's HRI architecture.

2 The project MOnarCH

The European project MOnarCH targets: (i) the development of a novel framework to model mixed human-robot societies, and (ii) its demonstration using a network of heterogeneous robots and sensors, in the pediatric area of an oncological hospital. The robots will perform edutainment activities and will deal with the uncertainties introduced by people and robots themselves in order to obtain natural interactions.

Key properties of networked robot systems, such as robustness and dependability, are a major concern of the project. In addition, guidelines to translate the MOnarCH system to applications in hospital environments, and other scenarios sharing similarities with them, e.g., kindergarten, and personal assistance to elderly at home, will be delivered.

Moreover, innovation is expected mainly in: (i) the modeling and analysis of the dynamics of social organizations, and social individuals, (ii) the mapping between such models and implementable systems, (iii) the integration between models related to social and asocial behaviors, (iv) the introduction of creative methods of interaction between people and robots, based on models of the dynamics of social organizations and individuals, and (v) the adaptation of robots to individuals and groups of people.

2.1 The MBots

The MOnarCH robots (MBots) (Fig. 1) have been designed to establish social interactions with the children who are hospitalized in the pediatrics ward of the hospital. The MBots are 1m height robots equipped with 4 omnidirectional mecanum/wheels; 2 laser range finders (LRF) for 360º coverage; and an RGB-D (Red, Green, Blue, - Depth) sensor for close obstacle detection. They perceive the environment with a second on-board RGB-D sensor; several touch sensors in their shell; a microphone; and an RFID reader. The robots also get video information from network of distributed omnidirectional cameras installed in the hospital. For interacting the MBots are equipped with several devices: Two 1-DOF arms; a 1-DOF Head (pan); two speakers to reproduce verbal and non-verbal utterances; a touch screen in the robot's chest that enables it to show images, videos, and menus to get input from the user (buttons, sliders, etc.); a video projector; several RGB LEDs installed in the robot's eyes, cheeks, and base; and a LED matrix that acts as the robot's mouth.

Fig. 1. An MBot Robot interacting with a child.
© Miguel A. Lopes / LUSA

2.2 The Hospital Scenario

The robots operate in three different areas of the pediatrics ward of the hospital (Fig. 2): part of the main corridor (Fig. 2 down-left), the pediatrics ward playroom (Fig. 2 down-centre), and the pediatrics ward classroom (Fig. 2 down-right). Within these areas the robot's behaviors vary so they can adapt to the different situations they might encounter. In order to do that, the MOnarCH project defines three storyboards, where each one applies to a specific area of the hospital. The first one is the Joyful Warden, where the robots will patrol and socially interact with the people they encounter in the corridor and the playroom. The second storyboard is the Interactive Game, where the robots engage with the children in a collaborative game. The third scenario is the School Teaching Assistant, where the robot acts as a teaching assistant in the pediatrics classroom by projecting videos and animations related to the class content.

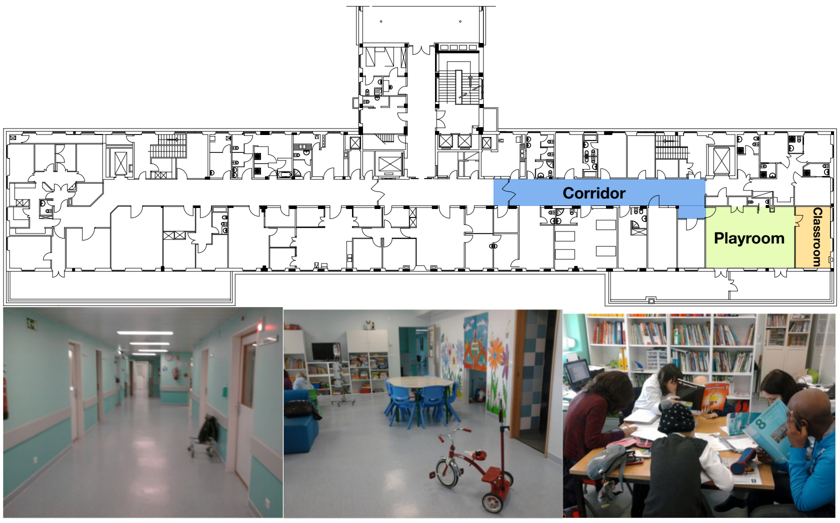

Fig. 2. The Hospital Scenario. The colored areas in the upper map correspond to the three scenarios in which the robots operate and move freely: corridor (down-left), playroom (down-centre), and classroom (down-right).

For some of these scenarios, we start developing the social interaction from previous work (Gonzalez-Pacheco, Ramey, Alonso-Martin, Castro-Gonzalez, & Salichs, 2011). However, in the MOnarCH scenarios, the robot's social behaviors must comply with different situations and social norms that might apply to each one of these scenarios. Also, the robots must be able to interact with children, hospital staff and bystanders. This means that the robots must behave in different manners depending on to whom they are interacting with, and on the current interaction context. In this way, we divided people in three different categories: hospital staff, children, and other people. We model these roles using the roles proposed in the literature (Scholtz, 2002) in the following way: hospital Staff will act as Supervisors and as Peers; children will act as Peers; and other people will be considered as Bystanders.

3 The HRI architecture for social robotics with children

The MOnarCH HRI architecture is depicted in Fig. 3. Its main components are the Dialogue Manager (DM) and the Multimodal Fusion (MFu) and Fission (MFi) modules. The DM receives the CA activations from the Socially Aware Planner (SAP). That is, the SAP[1] is aware of which is the current storyboard and situation and decides which CAs might be appropriate for such context.

The Dialogue Manager (DM) is the element of the architecture that manages the execution of the MOnarCH HRI. By managing we mean that it does not execute directly the interaction but, rather, it orchestrates the actions that compose this interaction by sending the appropriate commands to the robot's Multimodal Fission (MFi) modules, which are actually in charge of the execution. In short, we can say that the DM decides what to do while the MFi knows how to do it. For instance, if the robot needs to say a something to a user, the DM decides which sentence to say, with which emotion, volume etc. Then, it sends this command to MFi, who knows how to perform the utterances of that phrase. The interaction can be executed, not only by voice, but using other interfaces such as gestures, led expressions, images in the screen, etc. This multimodality applies also to the inputs: besides of listening to what the user says, the robot also uses other sensors such as touch, RFID readings, cameras, etc. The multimodality is supported thanks to the use of the Multimodal Fusion (MFu) for inputs and the Multimodal Fision (MFi) for outputs (see Fig. 3).

---

1 The functionality of the MOnarCH's SAP goes far beyond than telling the DM which CA might be needed. However, here we only describe the functionalities related to MOnarCH HRI.

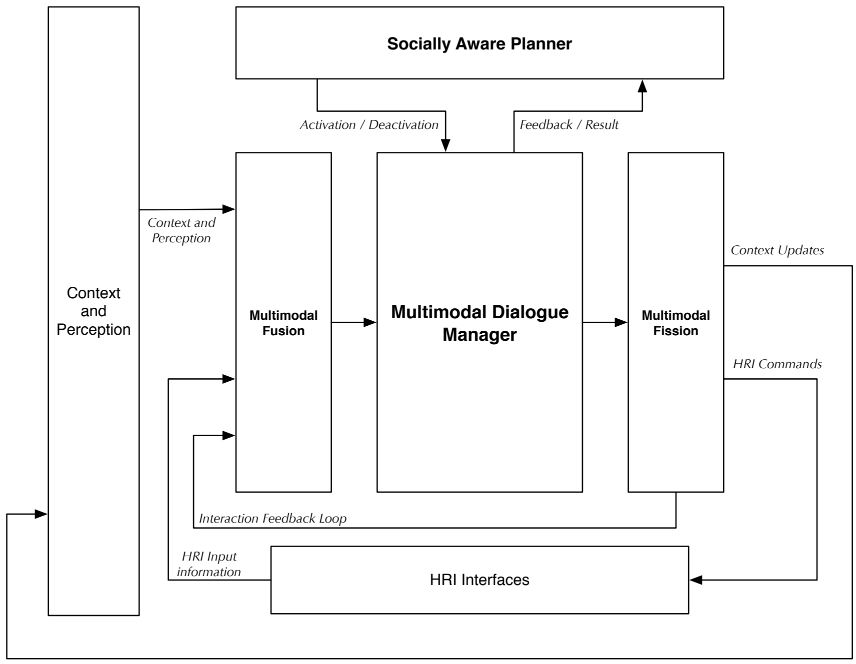

Fig. 3. Overview of the MOnarCH HRI architecture.

The MFu gathers the information coming from the input HRI interfaces and from the Perception, and second, aggregates and translates them into a format that is useful for the DM. With this information, the DM can decide what to do and send the appropriate commands to the MFi. The MFi is symmetric to the MFu but, in this case, is the DM who sends messages to the MFi, which is composed of translators that translate and forward the DM messages to the HRI interfaces. Using the two layers of translators (MFu and MFi) permits to keep the DM independent of the hardware that it is being used. For instance, it is possible to change the entire Automatic Speech Recognition system with only replacing the translator node of the MFu2.

The DM incorporates a rule-based production system inspired by the COLLAGEN plan trees (Rich, Sidner, & Lesh, 2001). The rules of the

2 However, note that the MFu is not only composed of translators. It also incorporates a layer of nodes that aggregate the information coming from the interfaces.

production system, which are called recipes or dialogues, enable managing the dialogues with aspects of finite-state, frame, and information-state-based approaches. These recipes are written in XML files, and define the preconditions that trigger the execution of a certain CA and the actions that should be executed in such CA. That is, these rules codify what to do when certain inputs are received. For instance, consider the case of a *give greetings* CA. Once this CA is activated by the Socially Aware Planner, we can write a precondition in the recipe which waits for the input from the perception indicating that a user is in front of the robot. Once this precondition is triggered, the recipe can contain different kind of actions. For example, to execute a greetings gesture, say a random greeting utterance, express a big smile, etc. Note that the MFu usually feeds the information that triggers the preconditions, while the actions to be executed are sent to the MFi. Once the recipe is ended, the results of the interaction are sent to the Socially Aware Planner.

4 Conclusions

We presented an HRI architecture that enables the robots developed in the MOnarCH project to establish social relations in different scenarios. The naturalness of these social interactions is quite important in the context of the project MOnaCH since the principal users are children who are hospitalized in the pediatrics ward of an oncological hospital. Since the main goal of these robots is to try to engage those children in edutainment activities, the HRI architecture presented in this work becomes one of the key points of the project.

Using the approach presented, the robot is able to communicate with the user in a very natural way thanks to the Multimodal Fusion (MFu), the Dialog Manager (DM), and the Multimodal Fission (MFi) structure at the core the HRI architecture.

The MFu module aggregates the information perceived and sends it to the MD that decides which message to transmit and then it communicates this message to the MFi. The MFi allows the robot to express the same communicative expression in different ways at different moments. Therefore, the resulting HRI architecture provides a natural, flexible, multimodal communication between the robots and the children.

Acknowledgments

The research leading to these results has received funding from several projects: the European Union Seventh Framework Programme FP7/2007-2013 - Challenge 2 - Cognitive Systems, Interaction, Robotics - under grant agreement No 601033 – MOnarCH; The project called: Desarrollo de robots sociales para ayuda a mayores con deterioro cognitivo / Develo-pment of social robots to help seniors with cognitive impairement (ROB-SEN) funded by the Ministerio de Economía y Competitividad (DPI2014-57684-R); and from the RoboCity2030-III-CM project (S2013/MIT-2748), funded by Programas de Actividades I+D en la Comunidad de Madrid and cofunded by Structural Funds of the EU.

References

Austin, J. L., 1975. How to do Things with Words. Oxford University Press.

Gonzalez-Pacheco, V., Ramey, A., Alonso-Martin, F., Castro-Gonzalez, A., & Salichs, M. A., 2011. Maggie: A Social Robot as a Gaming Platform. International Journal of Social Robotics, 3(4), 371–381. doi:10.1007/s12369-011-0109-8

Jakobson, R., 1960. Closing statements: Linguistics and Poetics. In T. A. Sebeok (Ed.), Style in language (pp. 350–377). Cambridge, MA, USA.

Rich, C., Sidner, C., & Lesh, N., 2001. Collagen: Applying Collaborative Discourse Theory to Human-Computer Interaction. AI Magazine.

Scholtz, J., 2002. Theory and evaluation of human robot interactions. In 36th Annual Hawaii International Conference on System Sciences, 2003. Proceedings of the (Vol. 3, pp. 125–134). IEEE. doi:10.1109/HICSS.2003.1174284

Searle, J. R., 1969. Speech Acts. Cambridge University Press.