Assessing Children-Robot Interaction in a Pediatrics Hospital Ward: The MOnarCH Case*

João Silva Sequeira

Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001 Lisbon, Portugal,

joao.silva.sequeira@tecnico.ulisboa.pt

Maria Isabel Aldinhas Ferreira

Centro de Filosofia da Universidade de Lisboa. Faculdade de Letras, Alameda da Universidade 1600-214. Lisbon. Portugal,

isabelferreira@letras.ulisboa.pt

Abstract. The present paper discusses the assessment of the ongoing experiments in children-robot interaction under project MOnarCH. The main concerns while planning the experiments and designing the evaluation metrics are discussed.

Keywords: Human-Robot Interaction (HRI), Metrics, Behavior segmentation.

1 Introduction

The MOnarCH Project[1] aims at developing a system of networked robots suited to interact with humans, especially with children in a pediatric hospital. Most of this interaction is planned to occur in the context of edutainment activities that include (i) helping to maintain the children in a socially interesting dynamics (ii) playing with children, and (iii) acting as school teaching assistants.

Aware of the specific problems that child-robot interaction poses, problems that go far beyond those that emerge in general HRI, the project defined as a first priority to identify the children's expectations in what concerns the roles/functions to be performed by a robot as well as the definition of its probable physical appearance.

* This work was supported by EU project FP7-ICT-2011-9-601033-MOnarCH and Portuguese FCT [UID/EEA/50009/2013].

1 Multi-Robot Cognitive Systems Operating in Hospital (www.monarch-fp7.eu).

Based on a survey that involved about 120 children and teenagers aged 7-16, an architecture was designed, taking into account not only the main functionalities referred by the children, in the survey, but also the anticipated physical appearance [1].

This description of a robot by the children involved in the survey, aimed at helping to design the architecture in a user-centered perspective enhancing, this way, the engaging potential and interactive possibilities of the interface.

2 Assessing Child-Robot Interaction

2.1 Preliminary Issues

In general, measuring the performance of a social robot in a real social environment is subject to multiple biasing factors. Often, seemingly innocuous actions may compromise the confidence in the results obtained.

To reduce the influence of such biasing factors one may resort to long-run experiments, where any transient effects may simply vanish in time.

Given that one of the goals of the project is to study the establishment of bidirectional relationships between humans and robots in a social context, short-term experiments may just sample the environment at favorable periods hence biasing the conclusions. Long duration experiments are thus the natural choice to assess the quality of the HRI, namely that between children and robots in the hospital context.

This is the strategy being followed in the MOnarCH project, where social robots are placed in the Pediatrics ward of an Oncological hospital to interact mainly with children in edutainment activities.

Roughly, assessing HRI in a social environment must account for the dynamics of the social environment. Such dynamics can be expressed in terms of events marking relevant time instants along the duration of the experiments. Hence, the number of different events and their frequency are important factors in HRI assessment. Multiple short-term experiments may be statistically equivalent to a single long-term experiment, provided that they capture all aspects of the dynamics of the environment. In a sense, the experimentation periods must be defined using a priori information on the dynamics of the environment.

2.2 Recording experiments, segmentation, and micro-behaviors

Classical analysis of HRI experiments outside the lab environment relies often in the segmentation of video sequences in micro-behaviors (see for instance [2]), which, in a sense, are delimited by the aforementioned relevant events. This is the approach being followed in MOnarCH.

Video naturally provides rich information on the experiment that can further be complemented with information from other sensors (static or placed onboard the robots). Metrics of interest include (i) the frequency of each micro-behavior, (ii) its duration, and (iii) its starting time in the sequence.

In general, segmentation of video sequences must be done manually, from recorded sequences, thus being an error prone process. Multiple long-term experiments are

required in order to obtain statistically valid information. The labor effort used at this stage tends to be significant.

The selection of the micro-behaviors to search for (in a sense the aforementioned classes of events) in the video sequences is of course an important step. The option in MOnarCH was to segment accounting for events that occur normally in the environment both in the presence and absence of robots (see Figure 1 and Figure 2 for examples). The current list can be checked in [3,4].

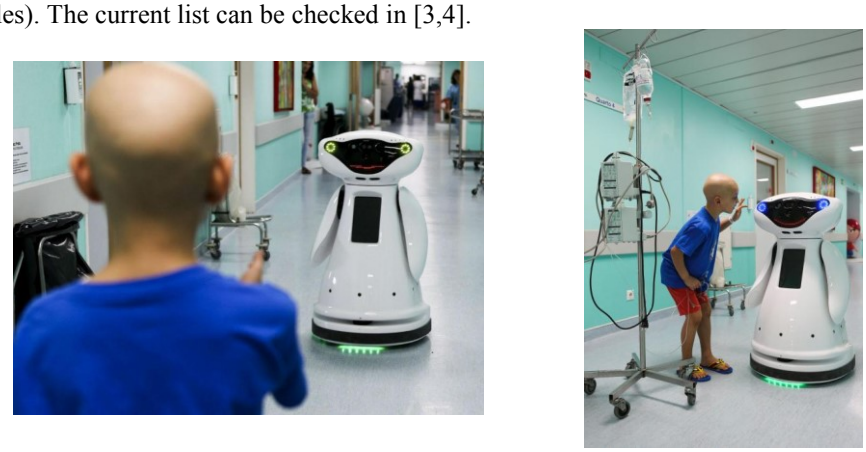

Figure 1: Interaction in close proximity

Figure 2: Interaction from a distance (left) and in proximity (right). These cameras are part of a larger set from where the video sequences are extracted.

3 Metrics

Following the ideas above, assessment metrics are based on the rate of activation of micro-behaviors of specific subsets. The hospital environment in MOnarCH dynamics alternates between fast and slow dynamics, depending mainly from the medi-

cal evolution of the inpatient children. Assuming the completeness of the list of micro-behaviors, activation rates capture the dynamics of the environment.

Social robotics acceptance depends on both objective and subjective factors. Objective factors are related mainly with robustness of the equipments/robots and safety of people. Subjective factors include the scientific/technology knowledge and, more generically, the cultural level of the inhabiting population.

Objective factors include, for example, the number and duration of failures during a specific period. These lead to widely known performance metrics.

In MOnarCH, subjective factors are estimated using the above idea of activation rates. Currently, three of such metrics are being used, measuring (i) the generic (positive and negative) influence of the robot in the environment, (ii) a positive influence on the inpatient children of having a robot in the environment, (iii) the effectiveness of the HRI interfaces onboard the robot, and (iv) assessing whether significant differences in the inhabitants between the environments with and without the robot can be detected. Details can be found in [3,4].

The valuation of each of the above metrics can be interpreted as representing an average activation rate of any micro-behavior in the metric. Decision thresholds can be constructed either using (i) the same or (ii) different sets of micro-behaviors as the metrics themselves. In MOnarCH, the second option is being used, meaning that there is a direct comparison between micro-behaviors relevant and non-relevant for the metric. The aim of the strategy is to increase the discrimination power of the metrics.

4 Conclusions

Current assessment is fully positive, i.e., (i) the acceptance of a robot by the children and people in general is confirmed, and (ii) the introduction of the robot in the hospital (social) environment was made without any relevant changes in the daily routine of the inhabitants. Numerical data can be found in [3,4]. The experiments are still ongoing aiming at assessing the effect of more complex HRI features.

5 References

1. Ferreira, MIA and Sequeira, JS. The concept of [robot] in children and teens: some guidelines to the design of social robots. International Journal of Signs and Semiotic Systems, 3(2), July-December 2014, special issue on “The Semiosis of Cognition”.
2. Dautenhahn, K.and Werry, I. “A Quantitative Technique for Analysing Robot-Human Interactions”. Procs. of the 2002 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, Switzerland, 2002
3. MOnarCH project Deliverable D7.7.1, Available from www.monarch-fp7.eu
4. Ferreira I.A. and Sequeira, J. (2015) “Assessing Human-Robot Interaction: The Role of Long-Run Experiments”. Procs. of the 18th International Conference on Climbing and Walking Robots and Support Technologies for Mobile Machines (CLAWAR 2015), HangZhou, China, 6-9 September 2015.