2510-2591 Joline Scheffler, Karola Pitsch (2020): Pre-Beginnings in Human-Robot Encounters. In: Proceedings of the 18th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing on the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2020_p02

Pre-Beginnings in Human-Robot Encounters: Dealing with time delay

Joline Scheffler & Karola Pitsch
Institute of Communication Studies, University of Duisburg-Essen
joline.scheffler@uni-due.de, karola.pitsch@uni-due.de

Abstract. Based on a video corpus and logfiles of HRI in a museum guide scenario the paper investigates a model of a stepwise process for opening a focused encounter. Towards developing an autonomous system, a Wizard of Oz set-up is used to investigate strategies for dealing with a system's time delay in the 'pre-beginning'. Initial Analysis shows that the wizard orients to the visitor's head orientation for proceeding to next steps.

1. Introduction

If we wanted to deploy robotic systems in everyday settings and collaborative tasks with human users, the technical systems need to be endowed with 'social' skills: users should be able to deal in an intuitive manner with the technical system making use of natural means of communication. To get in contact with a human and to open a focused encounter constitutes a central task and challenge for the robotic system: Not only do the potential participants have to signal to each other whether they are willing to engage in a focused encounter and what the nature of their exchange may be, but also they have to explore this in situ through fine-grained interactional coordination (e.g. Gehle et al., 2017). Particularly important is the moment before the first verbal greeting is exchanged (called 'pre-beginnings') during which the participants monitor each other with regard to the other's focus of (visual) attention,

body orientation and behavior in space. Therefore, we investigate such 'pre-beginnings' of 1:1 human-robot-interactions in the scenario of a museum guide robot. Towards modeling the opening strategy for an autonomous system, a laboratory study has been conducted with a humanoid robot (Nao, v4) making use of a human wizard to decide about the timing and choice of the robot's next actions.

Previous studies (Gehle et al., 2017) have pointed to interactional trouble caused by the robot's misdetection of faces and found that human wizards use creative strategies based on the robot's limited interactional resources to deal with them. In this paper, we focus on the phenomenon of time delay (2 seconds) occurring after the initial establishment of mutual gaze and investigate the following question: How do wizards attempt to deal with time delays within the 'pre-beginning' phase? – Analysis leads, in the long run, to considerations about pre-planned actions and the unpredictability and contingency of social interaction (Schegloff, 1996).

2. Design of the Robot's Opening Strategies

To explore ways of enabling a robot system to get in contact with a user and to engage in a focused encounter, an initial model for an opening strategy focusing on the 'pre-beginning' in a museum guide scenario has been developed (Fig. 1, see also Gehle et al., 2017). It uses a limited set of communicational resources based on the restricted possibilities of the current robot and was implemented as a finite state machine and linked to the robot's visual perception. In this study, it was realized as a Wizard-of-Oz approach: The human wizard assumed the role of deciding about the appropriate next move and its precise timing based on the robot's visual perception and the states in the model. This allows to gain information to further tailor the model for the autonomous system. It consists of the stages:

(a) **'Look around'**: To increase the robot's limited field of view (60°) to detect a potential interaction partner and to display availability, the robot's idle strategy was designed as a 'look around' behavior. The robot moved its head continuously from one side of the room to the other (160°), attentive to detect faces and to wait for the wizard's decision about *when* to look at the visitor.

(b) **'Look at'**: When the wizard – using the robot's camera view – decides to look at the visitor and presses the button on the GUI (Fig. 1a), the robot autonomously adjusts its head towards a detected face. As the face detection is realized by the robot's algorithms, misdetections might occur (s. also Gehle et al., 2017).

(c) **'Talk to' vs**. **'Look away'**: While the robot is in the state of looking at a (sometimes: presumably) detected face, the wizard has the choice to either make the robot produce a verbal greeting 'Hello, I am Nao' (i.e. 'talk to' on the GUI) or to make the robot lower its head orientation (i.e. 'look away' on the GUI) (Fig. 1c). Sometimes, these options appear with a delay of about 2 seconds – the GUI displaying "Anything to do" (Fig. 1b).

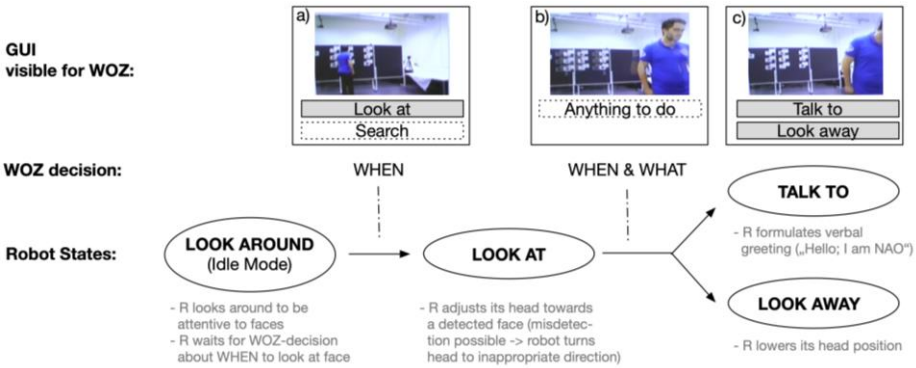

Figure 1. Model of opening strategy with corresponding GUI and decisions of the human Wizard.

3. Study, Data and Analytical Method

A Nao robot was set up to act as museum guide offering information about three areas of a museum exhibition which was recreated in the laboratory. The wizard (in total: 3 wizards) was placed behind a wall in the same room to have at his/her disposal (i) auditory information, and (ii) the robot's internal vision of scene which was transmitted online to the wizard's GUI with a delay of about 3 to 5 ms. The data corpus is comprised of 38 (34 complete) recordings from 4 external HD video cameras, the robot's internal VGA camera, logfiles of the dialogue states and recordings of the visitors' spatial conduct by an independent Kinect camera. In 10 cases, we also video-taped the wizard manipulating the mouse when pressing the buttons which allows insights e.g. into the wizard's *preparation* for pressing a button (analyzed here). Analysis is based on Conversation Analysis (EM/CA).

4. Case Analysis: Time delay in moment of transition

The case analysis (VP-31) will show that (i) a time delay occurs after the 'look at' stage due to technical reasons. (ii) This is dealt with by the wizard through a constant orientation to the features 'mutual gaze' and 'head orientation in face-to-face position' for passing to the next stage in the model.

(a) Establishing mutual gaze with visitor: The visitor V enters the room (Fig. 2a). The wizard's (W) GUI shows the option 'look at'. At #00:01:44 W presses 'look at'. V's body and head orientation is manly directed toward exhibits. W starts the pre-beginning when mutual gaze between Nao and V is achieved (Fig. 2b).

(b) System's time delay after wizard choice of 'look at' leads to visitor reorienting to exhibit: After pressing 'look at' Nao's system causes a two second time delay, visible on the GUI as 'anything to do'. With the lack of a verbal utterance produced by the robot, the visitor smiles and turns his upper body to the nearest exhibit (Fig. 1, Fig. 2c).

(c) Wizard anticipating visitor to reappear in face-to-face-position: The system's time delay is also an *interaction time delay*, i.e. to verbally continue. W waits with the cursor on the 'talk to'-button to greet V (Fig. 2d). W anticipates V to show availability to Nao at the transition from one exhibit to the next or Nao.

(d) Wizard pressing button 'talk to' when visitor reappears face-to-face to the robot: In transition from exhibit to Nao, V reappears in face-to-face position. W presses 'talk to' at #00:01:53, reacting to the mutual gaze (Fig. 2e).

Figure 2: Timeline of wizard's and visitor's activities building up to pre-beginning and opening

5. Conclusion

The opening model suggests a stepwise process in which the succession of the stages 'look at' and 'talk to' might lead to a technology-based time delay. This causes a particular opening situation of 'time stretching' and requires further forms of micro-coordination (e.g. to deal with unexpectedness). Analysis has shown that the wizard strongly orients to the visitor's gaze direction and head orientation for activating next steps. Future work will extend analysis and design considerations.

6. Acknowledgement

The authors acknowledge the financial support from Volkswagen Foundation (Dilthey Fellowship) and 'CITEC' (EXC 277, IP-18), Bielefeld University.

References

Gehle, R, Pitsch, K, Dankert, T. and Wrede, S (2017): How to Open an Interaction Between Robot and Visitor? Strategies to Establish a Focused Encounter in HRI. *HRI* 2017, pp. 187-195.

Schegloff E. A. (1996): Issues of Relevance for Discourse Analysis: Contingency in Action, Interaction, and Co-Participant Context, in E. H. Hovy and D. R. Scott (eds.): *Computational and Conversational Discourse: Burning Issues - An Interdisciplinary Account,* Heidelberg, pp. 3-38.