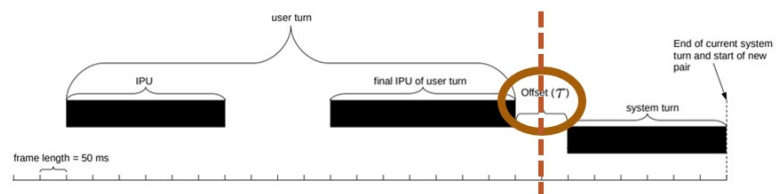
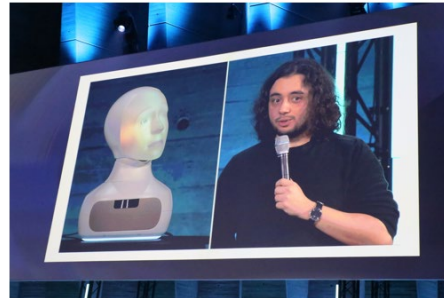


Master's Project

Modeling and Evaluating Rules of Turn-Takings between Humans and Robots

Description

In human-robot spoken interaction, an important feature to consider is when it is appropriate for the robot to speak [1]. Indeed, turn-taking is an essential mechanism in language interactions. Therefore, identifying and predicting when the robot must hold or shift the floor improves the acceptability of conversational robots by humans. A state-of-the-art, yet naïve approach, is to make turn-taking predictions at silences [2-3]: taking the turn should be prompt when a given agent is explicitly addressed, but silences should not be too short, to avoid the sensation of being interrupted; similarly, pauses should not be too long, which can also feel unnatural [4]. Thus, turn-taking predictions depend then on the temporal occurrence of pauses, the duration of which specifies whether an agent is about to shift the floor or to hold it.



This project will focus on modeling rules of human-robot turn-takings. The focus will be on learning white-box models, that is, models that are explainable and understandable by system designers (as opposed to black-box models, e.g. neural networks [5]). This can be done by means of set of rules, or specifications of the form "between 250ms and 1s, robot should not speak". For instance, recent data-based approaches [6-7] extract interpretable formalisms, that have this kind of resemblance with natural language, to depict complex human-robot interactions, including turn-takings [8].

Goals

Your task will be to create a user study, where participants will converse with a robot (e.g., a Furhat). You will design when the robot ideally has to take the turn, run experiments, and readjust your turn-taking model regarding the participant's feeling of whether the robot took the turn appropriately.

Requirements

The project is suitable for students interested in the fields of robotics, human-robot interaction, the design of user studies and machine learning. A good command of Python is required.

Contacts

Alexis Linard (linard@kth.se), Ilaria Torre (ilariat@kth.se)

References

[1] Harvey Sacks, Emanuel A Schegloff, and Gail Jefferson. 1978. A simplest systematics for the organization of turn taking for conversation. In *Studies in the organization of conversational interaction*. Elsevier, 7–55

- [2] Luciana Ferrer, Elizabeth Shriberg, and Andreas Stolcke. 2002. Is the speaker done yet? Faster and more accurate end-of-utterance detection using prosody. In seventh international conference on spoken language processing.
- [3] Antoine Raux and Maxine Eskenazi. 2009. A finite-state turn-taking model for spoken dialog systems. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*. 629–637.
- [4] Iolanda Leite, Hannaneh Hajishirzi, Sean Andrist, and Jill Lehman. 2013. Managing chaos: models of turn-taking in character-multichild interactions. In *Proceedings of the 15th ACM on International conference on multimodal interaction*. 43–50.
- [5] Gabriel Skantze. 2017. Towards a general, continuous model of turn-taking in spoken dialogue using LSTM recurrent neural networks. In *Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue*. 220–230.
- [6] O. Maler and D. Nickovic, “Monitoring temporal properties of continuous signals,” in *Formal Techniques, Modelling and Analysis of Timed and Fault-Tolerant Systems*. Springer, 2004, p. 152–166.
- [7] G. Bombara, C.-I. Vasile, F. Penedo, H. Yasuoka, and C. Belta, “A decision tree approach to data classification using signal temporal logic,” in *Proceedings of the 19th International Conference on Hybrid Systems: Computation and Control*. ACM, 2016, pp. 1–10.
- [8] David Porfirio, Allison Sauppé, Aws Albarghouthi, and Bilge Mutlu. 2018. Authoring and verifying human-robot interactions. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*. 75–86.