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


