Master's Thesis Proposal

Adapting robot's explanation for failures according to observed human behavior

Description

The objective of this thesis is to adapt the robot failure explanation by using multimodal human behavioural changes after observing the failure. We have a multimodal dataset of previously recorded data of human reactions (behaviour change, ΔB) to robotic failures and robot explanations [1]. The dataset is used to analyse ΔB across different behaviour metrics. The aim is to develop an algorithm to map ΔB to different explanation levels by learning a representation of ΔB using suitable machine learning methods. Finally. the algorithm is tested online in a human-robot collaborative experiment involving pre-planned robotic failures, where the robot should adapt its explanation level based on observed ΔB.

When a robotic failure occurs in a human-robot collaborative task, the robot can ask for human help in the form of resolution action by giving an explanation. We divide the robotic explanation into two parts: explaining the cause of failure and explaining the resolution action. The content of the explanation is dependent on multiple factors, such as the severity of failure, human understanding of the robotic system, previous interaction with the robot, etc. In a user study [1], we recruited 68 participants to interact with a robot that failed at different stages of executing a task and provided different levels of explanations to resolve the failure and finish the task collaboratively with the human. We recorded several aspects of the interaction: video recordings of the participant's face, human body pose tracking, and audio recording. We aim to use this multimodal data to understand the link between human behaviour change and the appropriate level of explanation that leads to successfully completing the task. More details regarding the study can be found in [1][2].

Goals

You will start by extracting information from the recordings of the human-robot interaction from a previous study. The recording for each participant contains a video recording of the participant (from the robot's POV), a video recording of the interaction (an overhead camera), human body pose tracking, and an audio recording (participant's verbal statements, reactions, commands, and requests). Thus you will create a dataset of human reactions, i.e., behaviour change (ΔB) to robot failures followed by the explanations given by the robot. Further, you will analyse ΔB across different Behavior Metrics:

a. Face,
b. Bodypose,
c. Gaze,
d. Audio (Verbal statements/commands/requests), and more.

You will analyse the dataset to find trends in ΔB for different failures. different explanation levels and different strategies for varying explanation levels. You will also investigate which metric represents ΔB better: a, b, c, d or a combination of these? You will utilise suitable machine learning techniques to form a

representation for the observed ΔB. Further, you will create an algorithm to map different explanation levels to the observed ΔB, which would allow the robot to adapt its explanation level. You will further test this algorithm in a human-robot experimental study with pre-planned robotic failures and explanations. It is also possible to use reinforcement learning techniques to change the explanation level according to ΔB observed in repeated interactions with the same participant.

Requirements

This project is apt for students looking to apply ML methods to practical problems.
Proficiency in Python and familiarity with ML techniques is required.
Some experience with robots and ROS is a plus.

Contacts

Parag Khanna (paragk@kth.se) Elmira Yadollahi (elmiray@kth.se)

References

[1] P Khanna, E Yadollahi, M Björkman, I Leite, C Smith, “User Study Exploring the Role of Explanation of Failures by Robots in Human Robot Collaboration Tasks”. in HRI 2023: An interdisciplinary Workshop on the role of failure in HRI, arXiv preprint arXiv:2303.16010
[2] P Khanna, E Yadollahi, M Björkman, I Leite, C Smith, “Effects of Explanation Strategies to Resolve Failures in Human-Robot Collaboration”. IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN) 2023. arXiv preprint arXiv:2309.10127