Abstract—Spoken language can be an efficient way to warn robots about threats. Guidance and warnings from a human can be used to inform and modulate a robot’s actions. An open research question is how the instructions and warnings can be integrated in the planning of the robot to improve safety. Our goal is to train a Deep Reinforcement Learning (DRL) agent to determine the intention of a given spoken instruction, especially in a domestic task, and generate a high-level sequence of actions to fulfill the given instruction. The DRL agent will combine vision and language to create a multi-modal state representation of the environment. We will also focus on how warnings can be used to shape the DRL’s reward, concentrating on the recognition of the emotional state of the human in an interaction with the robot. Finally, we will use language instructions to determine a safe operational space for the robot.

I. INTRODUCTION

In the future, robots are expected to work as companions with humans in various areas including domestic scenarios such as care-giving. Human-robot interaction safety has not been well studied [1]. Even with well-engineered robots, it would be unrealistic to move robots directly from factories to home environments to perform complex tasks [2] due to safety [3]. Moreover, robots also have to continuously adapt to new environments to avoid hazardous actions since using experts to program a robot for every environment is impossible. Hence, we need adaptive learning algorithms.

Spoken language can be considered one of the most effective communication channels to warn robots about threats. For example, robots may not notice an external threat or mis-planning that may harm a human or the robot itself. However, a human can warn or guide the robot by a verbal utterance toward a safer interaction. How robots react to safety warnings is not addressed exhaustively in the literature. The closest related research area is assigning tasks to robots by verbal instructions [4]. They follow rule-based methods to utilize spoken language instructions which can cover only a limited number of scenarios.

Our goal is to train a robot to safely perform complex tasks with the ability of processing environmental feedback, including guidance and warnings by a human, to shape a proper signal for updating its own policy. Therefore, our research is focused on three capabilities of the robot: generating high-level actions from verbal instructions, extracting reward from prosodic/sentiment features of the human speaker, and learning a safe workspace for the robot.

*This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 642667 (SECURE).

II. FOCUS AREAS

A. Mapping Spoken Instruction to a Sequence of Actions

We introduced a framework to obtain the intention of a given spoken instruction (e.g. "boil water") and generate the sequence of actions ("moveto kettle", "grasp kettle", ...) to fulfill the task [5], [6]. We developed a symbolics environment from the “Tell Me Dave” Corpus [7] to train the RL agent. A distributed symbolic state representation reduced the learning time (about 4x) (e.g. {On Kettle Sink}, {Near Robot Sink}, ...) of the environment. In our case, the environment state was directly accessible through the simulation while this needs to be extracted in a real life scenario. Therefore, we will extend by encoding vision and instruction in a fused state similar to Shu et al. [8] in a more realistic simulator like Ai2Thor [9] (see figure 1).

![Fig. 1. The modular approach using intention detection and reinforcement learning trained for each objective to generate the sequence of actions[6].](image)

B. Extracting Reward from the Human Speech

The robot needs to continuously process human speech to detect implicit interruptions or any change in the instruction. The robot is expected to be able to stop (both soft and emergency) with a minimum latency in an unsafe situation (see figure 2). We developed a reinforcement learning approach to optimize the accuracy and latency concurrently [10]. As a result, our model achieved about 50% latency reduction with the same level of accuracy evaluated on the iCub recorded data in our lab. We also improved the robustness of emotion recognition by proposing data augmentation techniques like overlaying background noise [11]. As future work, emotion recognition will be used to filter warnings and to record this experience in the RL’s memory for updating the agent’s policy. We will use a pretrained model in simulation to focus on learning new safety cases in the real scenario.

C. Safe Human-Robot Collaboration in Manual Tasks

Safety becomes more important when humans work with robots collaboratively. For shaping such a collaborative scenario incrementally, as an initial step, we improved the learning of the Deep Deterministic Policy Gradient (DDPG)
are used in different areas and we focused on the high-end-to-end approach.

As a next step, we will concentrate on obtaining state representations in real life scenarios. In parallel, we proposed a model to detect angry emotions rapidly, which can be used as an implicit interruption to planning to lead to a safer human-robot interaction. As future work, we will focus on how the robot can learn from experience to immediately avoid the same behavior. We also investigate how teaching the operational space to the robot can be performed intuitively. We plan to extend this to a small kitchen scenario which can bring together all these ways of using spoken instructions/warnings to guide the robot towards safer interaction.

REFERENCES


