Mediating Joint Intention with a Dialogue Management System

Anonymous Author(s)

ABSTRACT

A necessary skill enabling robots to take part in decision making processes with their human user is the ability to mediate the joint intentions. This paper proposes an initial formulation of an architecture to mediate joint intentions between a social robot and a human participant. The architecture is based on a combination of plan recognition techniques to identify the user intention, and a Reinforcement Learning network which learns how to best interact with the inferred intention. This research is in early stages and here we report its current progress.

KEYWORDS

Joint Intention, Robotics, Goal Recognition, Reinforcement Learning

1 INTRODUCTION

The socio-technological evolution of human society has motivated the integration of robots in social and personal spaces. Hence, it is becoming a pressuring requirement for social robotics to understand human intentions and adapt to social values and needs.

Among other reasons, humans interact to understand and mediate intentions with other human participants [7]. A successful mediation of intention enable participants to decide their next contribution, to manage expectations, or to decide whether to trust the other participant. Natural language dialogues are among the primitive modes [1] of human-human interaction, and are also consistently used to mediate intentions.

Dialogue management strategies have exploited joint intention theory for building team dialogues [6]. However, this work views joint intention from a perspective of planning shared tasks [5] for a human and a robot participant. The objective of this work is to model *joint intention theory* for Human-Robot Interaction (HRI) in a household scenario. Within the scenario, we explore the cases where a person could need assistance from a robot such as: in cooking, finding different objects in the house, preparing for a visit to the supermarket, doctor or a friend. For instance, the person might say "I want to prepare a salad" to a robot, possibly having an intention for the robot to help her in cooking the dinner. However, its challenging for a robot to come to a common understanding of cooperatively performing the given task.

We define the objective of this research with the question: **how to create joint intention with machines?** We attempt to answer to the research question by proposing an architecture to mediate joint intentions. The proposed architecture offers a turn-based interaction scheme that allows two participant (human and a robot) to mediate an intention regarding a shared task.

2 METHODOLOGY

We formally define an intention as a plan $\pi = \{a_0, a_1, ..., a_n\}$ together with a goal g an agent is committed to [7]. π can either be a complete plan achieving g, or a partial plan directed towards it. A joint intention is an intention that is shared by multiple agents, and where the plan describes what the agents are jointly committed to do while sharing the same goal. Therefore, a joint intention is a plan $\pi = \{a_0, b_1, a_2, ..., a_n\}$ together with a goal g. a and b signify that the actions in π can be allocated to any participant.

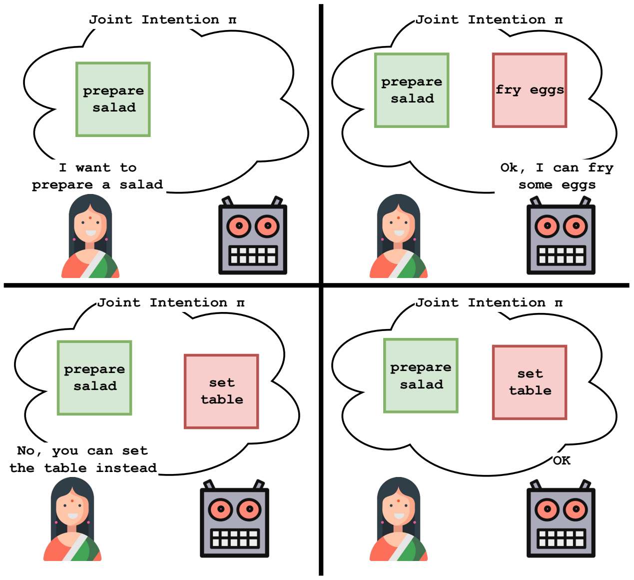

Figure 1: Creation of a joint intention with a robot. During its turn each participant adds and removes tasks (or primitive actions) from the shared intention π. Participants specify what they or the other will do until a common agreement is met.

Following the given definition, we propose an interaction mechanism that allows two participants to collaboratively build π, by being able to add or remove actions belonging to it (Figure 1). Currently, we have the following assumptions: first, for every trial two participants are present, that is a user (human) and the robot (dialogue manager (DM)). Second, the DM is modelled to be user initiated, which always proposes the first action that will enter the set π.

At every turn, the machine agent must evaluate π and decide whether to remove an action from it, or to add one. Formally, it has to carry the following operations:

1. Observe π and compute $P(G|\pi)P(\pi)$, the probability distribution over the possible goals i.e. perform *goal recognition*.
2. Compute for every possible goal $g \in G$, a plan π_g with the constraint $\pi \subseteq \pi_g$.
3. Compute the set of proposable actions $I = \bigcup_g \pi_g \setminus \pi$ and the set of removable action $R \subseteq \pi$.
4. Select an action from I or R, or do nothing. The selection of the action to perform should be based on $P(G|\pi)P(\pi)$.

This corresponds in specifying a *policy* $P(X|\pi)$, where $X = I \cup R \cup \{none\}$.

The following sections will be devoted to explain how we propose to implement points 1-4, by a combination of classical planning and reinforcement learning. Additionally, we propose how a speech interface for the system can be developed.

2.1 Goal Recognition and plan generation

For our initial experiment we implemented goal recognition and planning (points 1-3) using the Planning Domain Definition Language (PDDL) [2]. PDDL belongs to the group of planning techniques known as classical planning, and allows to easily create non-hierarchical task domains.

For a given task domain we select the set of goals G as possible goals the user can pursue. Goal recognition and plan generation is achieved by a modified version of the plan recognition method proposed in [4]. Our modifications allow the PDDL planner to plan using partially instantiated actions[1]. This allows the user to give partial specifications of the tasks he intends to do. Given a joint intention π, goal recognition is performed as:

$$P(G = g|\pi) = \alpha e^{C(\emptyset,g)-C(\pi,g)} P(G) \tag{1}$$

where $C(\pi, g)$ is the cost of a plan achieving g and constrained to contain π, $C(\emptyset, g)$ is the cost of an optimal plan achieving g without any constraint α is the normalizing factor. Hence, $C(\emptyset, g) - C(\pi, g)$ gives indication on how costly it is to deviate from an optimal plan achieving g for compliance with the intention π.

For every goal g we generate as π_g the optimal plan achieving g while being constrained to contain the joint intention π.

At the end of the planning phase, we obtain the set of actions the agent can propose to add to the joint intention $I = \bigcup_g \pi_g \setminus \pi$, together with the actions that can instead be removed $R = \pi$.

2.2 Learning the agent strategy with Reinforcement Learning

We propose to learn the system policy $P(X|\pi)$, $X = I \cup R \cup \{none\}$ (point 4) using a reinforcement learning (RL) algorithm. At every turn, a Q-Network [8] evaluates the current state of π together with the available actions X, selecting which actions $x \in X$ to perform by an ϵ-greedy policy computed on the expected return of the actions. In RL, agents learn which policy to adopt by maximising the reward they receive during each episode. The current version of the reward function is:

$$R = -\alpha T + \beta \frac{\overline{\pi} \cap \pi}{\overline{\pi} \cup \pi} + \gamma (C(\overline{\pi}, \overline{g}) - C(\pi, \overline{g})) \tag{2}$$

where $\overline{\pi}$ and $\overline{g}$ form the user's original intention (held when initiating the interaction). The first terms penalises every turn that the interaction takes, hence making interactions as short as possible. The second term evaluates how the final resulting intention is similar to the one the user had as objective for the interaction. The third term evaluates instead the cost the final mediated intention has, compared to the user's original one. α, β and γ determine how the three components of the reward function are weighted. Notice that the system cannot access $\overline{\pi}$ and $\overline{g}$, that are instead only used at the end of every interaction for evaluation. The goal of RL is thus to improve the unobservable intention $\overline{\pi}, \overline{g}$.

3 SPEECH INTERFACE

We further define two components to allow communication with the previously defined system using speech: a natural language understanding (NLU) component maps utterances to manipulations of the joint intention. And a natural language generation component (NLG) allows the robot to communicate the system action that it intends towards building the joint intention.

Inside the NLU component every sentence uttered by the user is:

(1) Classified into the system actions *add* or *remove*.
(2) The semantic roles of the sentence are classified into the objects and actions belonging to the task space.

This would allow to map a sentence such as "I want to prepare a salad" to the label *add*. The semantic roles *agent*: I, *verb*: prepare, *patient*: salad, could be used to instantiate an action **(cook user salad)** belonging to the task space and therefore utilizable inside π.

Symmetrically, the Speech Generation component should generate, from pairs $\{add, remove\} \cup (a \in A)$, a sentence expressing that system action. For example, the action *add*∪**(cook robot eggs)** could result in the utterance "I will fry some eggs".

4 FUTURE WORK

The research is still in its early stages and we are currently implementing the described system. We developed the goal recognition and the reinforcement learning agent components together with a simple user simulator. The user simulator is based on PDDL and simulates how the user would modify the joint intention during its turn, while following a randomly generated desired final intention $\overline{\pi}$ and goal $\overline{g}$. We use the simulator to train the Q-Network.

Initial experiments gave positive results, in the sense that the Q-Network being utilised is able to learn the structure of the problem for simple scenarios, and successfully maximises the possible rewards. Several investigations are needed and are still open to: what is the Q-Network learning? Does our current setting allows any generalisation? The current implementation requires hundreds of episodes to converge. Can the process be made faster/simpler? How to facilitate the online adaptation over real users?

Encapsulation of the joint intention model into a speech framework is still to be implemented. For early prototypes of the system we plan to implement the speech interface as described in Section 3. Later versions could see the implementation of a more complete SDS through for example a POMDP model [3]. This could allow to have dialogues that are not strictly related to the mediation of the joint intention, but rather more flexible and intuitive for the user. Investigation about properties and requirements that dialogues for mediating joint intentions should have is also to be performed.

REFERENCES

[1] New World Encyclopedia. 2017. Dialogue — New World Encyclopedia,. //www.newworldencyclopedia.org/p/index.php?title=Dialogue&oldid=1007366 [Online; accessed 24-January-2020].

[2] Drew McDermott. 2003. The Formal Semantics of Processes in PDDL. In *Proc. ICAPS Workshop on PDDL*. ICAPS Workshop, Trento, Italy, 8.

---

[1] We define a PDDL action as partially instantiated if not all of its arguments are bounded. An action is fully instantiated when all arguments are bounded.

[3] Michael Frederick McTear, Zoraida Callejas, and David Griol. 2016. *The conversational interface*. Vol. 6. Springer, Spain.

[4] Miguel Ramírez and Hector Geffner. 2010. Probabilistic plan recognition using off-the-shelf classical planners. In *Twenty-Fourth AAAI Conference on Artificial Intelligence*. MIT Press, USA, 6.

[5] David P. Schweikard and Hans Bernhard Schmid. 2013. Collective Intentionality. In *The Stanford Encyclopedia of Philosophy* (summer 2013 ed.), Edward N. Zalta (Ed.). Metaphysics Research Lab, Stanford University, USA.

[6] Rajah Subramanian, Sanjeev Kumar, and Philip Cohen. 2006. Integrating Joint Intention Theory, Belief Reasoning, and Communicative Action for Generating Team-Oriented Dialogue.. In *Proceedings of the National Conference on Artificial Intelligence*, Vol. 21. MIT Press, Boston, USA, 6.

[7] Michael Tomasello, Malinda Carpenter, Josep Call, Tanya Behne, and Henrike Moll. 2005. Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences* 28, 5 (2005), 675–691.

[8] Hado Van Hasselt, Arthur Guez, and David Silver. 2016. Deep reinforcement learning with double q-learning. In *Thirtieth AAAI conference on artificial intelligence*. MIT Press, Arizona, USA, 7.