Mediating Joint Intentions with a Dialogue Management System

Michele Persiani
Umeå University
Umeå, Sweden
michelep@cs.umu.se

Maitreyee Tewari
Umeå University
Umeå, Sweden
maittewa@cs.umu.se

ABSTRACT

A necessary skill which enables machines to take part in decision making processes with their users is the ability to participate in the mediation of joint intentions. This paper presents a formalisation of an architecture to create and mediate joint intentions with an artificial agent. The proposed system is loosely based on the framework of *we-intentions* and embodied 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.

KEYWORDS

Joint Intentions, 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 [16]. A successful mediation of intention enable participants to decide profitable collaboration, to manage expectations, or to decide whether to trust the other participant. Natural language dialogues are among the primitive modes [4] of human-human interaction, and are also consistently used to mediate intentions. Dialogue management strategies have exploited joint intention theory for building team dialogues [15]. However, this work views joint intentions with an accent on joint task planning [13] for a human and a robot participant, rather than on the communicative protocols being involved.

The objective of this work is to model mediation of intentions for Human-Robot Interaction (HRI) in a household scenario, and is loosely based on the framework of *we-intentions* [17]. 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.

Hence, we explore the following research question: **how to create joint intention with machines?** The motivation behind this research question belongs to desired specifications of AI systems, including the need of an integrated cognition and collaboration mechanism, and a natural interaction between human and AI systems. Ultimately, it is about investigating the boundaries between the Eco-system of AI with that of human-beings.

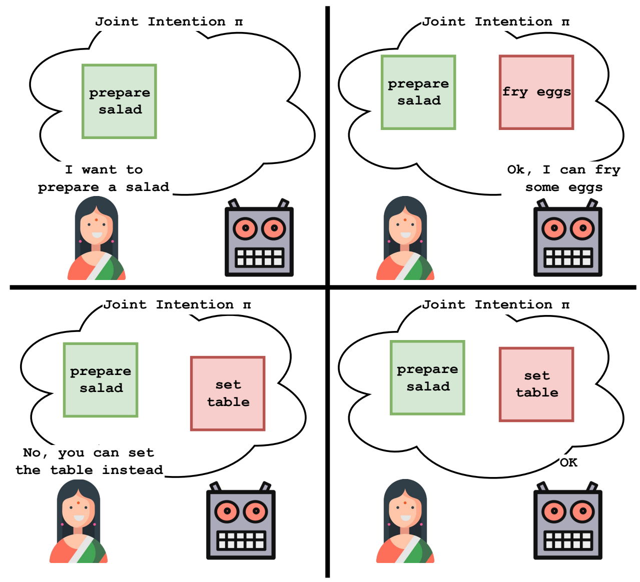

Figure 1: Creation of a joint intention with a robot. During its turn each participant adds and removes tasks (or primitive actions) from the shared intention π. Participants specify what they or the other will do until a common agreement is met.

To answer the research question, the formalisation of joint intentions in the context of shared task planning, and defining dialogue act functions [6] was done. This formalism offered a turn-based interaction scheme that allows two participant (human and a robot) to mediate an intention regarding a shared task.

2 METHODOLOGY

Some of the previous work [7, 11] proposed team rationality for building collaborative multi-agent systems, for example, in [11] the authors used Shared Plan [5] and *Propose Trees* to model collaboration as multi-agent planning problem, where a rational team will perform an action only if the benefits from performing an action is less than its cost. In [2] the authors formalised communication protocols using joint intention theory. The authors used joint persistent goals and persistent weak achievement goals to build joint intentions, and speech acts such as *request, offer, inform, confirm, refuse, acknowledge, and standing-offer* for their mediation.

As later described, we propose certain assumptions to lift some of the complexity that previous research utilizes in the—context of joint intention theory. We believe that such complexities, while theoretically sound, make implementations on real systems difficult and brittle; for this reason, we utilize a simplification of previous work's formalization for our needs. The rest of the section provides our simplified formalisation of mediating joint intention theory and attempts to briefly reason about the constraints posed.

Our proposed approach is based on predicate logic combined with planning, and is influenced by logic based semantics proposed in [1, 2, 14]. Agents are represented by $x, y, \dots x_1, x_2, \dots y_1, y_2, \dots$ and their actions by a_1, a_2, a_n. An intention of a single agent x is a plan $\pi = \{a_0, a_1, ..., a_n\}$ of actions together with a goal g the agent is committed to [16] and the intention is partially observed through $O \subseteq \pi$.

$Know(x, p) \equiv p \wedge Bel(x, p)$ represents the knowledge of agents and $MutBel(x, y, p)$ that x and y share a mutual belief about p—In our formulation an agent's intention is represented by the predicate $Intend(x, g, O)$ while a joint intention $JointIntend(x, y, g, O)$. An agent has an intention if following holds:

$$\begin{aligned} Intend(x, g, O) \equiv Know(x, \exists \pi\ O \subseteq \pi \wedge \\ Goal(\pi) = g \wedge \\ Commit(x, \pi)) \end{aligned} \tag{1}$$

i.e. not only is true that the agent has an intention and is committed to it, but the agent also has a belief about it. The set O is an explicit subset of π for which it is known that the agent already committed to it, and contains past observations or declarations about future commitments about π.

Eq. 2 expresses that to provide agents an intention doesn't require to explicit their full intention π, but only a part of it (see Figure 1), with the full intention being instead *inferred* by grounding the observed commitments in the task space.

A joint intention is an intention shared by the agents x and y with the same goal g. Therefore, a joint intention is a plan $\pi = \{a_{x0}, a_{y0}, a_{x1}, ..., a_{yn}\}$ together with a goal g where the actions in π can be allocated to either participants x or y. Furthermore, the involved agents have a mutual believe $MutBel$ about each others' commitments. Hence, two agents hold a joint intention if the following holds true:

$$\begin{aligned} JointIntend(x, y, g, O) \equiv Intend(x, g, O) \wedge Intend(y, g, O) \wedge \\ MutBel(x, y, JointIntend(x, y, g, O)) \end{aligned} \tag{2}$$

By this formulation x and y are allowed to have separate beliefs and inference mechanisms through which they find π; but are bound to have the same goal and observed commitments. Notice that this is a simplification of how joint intentions have been previously formalized in the literature, and to which we invite the reader. Nevertheless, this formalization is sufficient for our purpose of creating a dialogue manager that allows mediation of joint intentions. In this context of a dialogue between—two agents **x** and **y** we further make the following assumption:

$$\models \exists O \exists g\ JointIntend(\mathbf{x}, \mathbf{y}, g, O) \tag{3}$$

that translates as *there is always a joint intention between* $\boldsymbol{x}$ *and* $\boldsymbol{y}$. This assumption, while being quite strong, is quite reasonable for our context as the proposed DM is specifically tailored to mediate joint intentions. During every dialogue a joint intention is always obtained, and when the user leaves the conversation there is always an intention that was formed and is shared with the DM. Furthermore, it is always the case that the user utilises the DM to instantiate joint intentions. We do not take into consideration the cases in which the joint intention is bootstrapped or terminated as for example shown in [1].

Following the given definitions, we propose an interaction mechanism that allows two participants to collaboratively build O, by being able to add or remove actions from it. Currently, we have the following assumptions: 1) for every trial two participants are present, that is a human user and a Dialogue Manager (DM), that can be integrated for example in a house robot. 2) the DM is modelled to be user initiated, which always proposes the first action that will enter the set O.

Having an observed set O in a form of a partial plan, the DM can infer the most likely full intention π by utilizing plan recognition techniques as later described. This inference is based on the current state of the world that we assume to be available to the DM in the form of truth predicates. A possible architecture for maintaining an updated world description is not provided by this paper but can be for example implemented as in [3].

2.1 Goal Recognition and plan generation

At every turn of the dialogue the agent is required to infer the joint plan π to be able to participate in its mediation. For this purpose, we utilize plan recognition techniques based on the Planning Domain Definition Language (PDDL) [8]. 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 a user can pursue. Example of possible goals for Figure 1 can be to prepare dinner, lunch or breakfast. Plan recognition is achieved by a modified version of the method proposed in [10] with the following differences: 1) we allow the PDDL planner to plan using partially instantiated actions[1], and 2) the observations O are treated as a set rather than a sequence. Given an eventually empty set of observations O, goal recognition is performed as:

$$\hat{g} = \underset{g \in G}{argmax} \frac{C(\emptyset, g)}{C(O, g)} \tag{4}$$

where $C(O, g)$ is the cost of a plan achieving g and constrained to contain O, $C(\emptyset, g)$ is the cost of an optimal plan achieving g without being constrained by O. Hence, $0 \le \frac{C(\emptyset, g)}{C(O, g)} \le 1$ gives indication on how costly it is to deviate from an optimal plan achieving g for compliance with O. Finally, for an inferred goal $\hat{g}$ we obtain π as the optimal plan achieving g while being constrained to contain the observations O.

[1]We define a PDDL action as partially instantiated if not all of its arguments are grounded in the task domain. An action is fully instantiated when all arguments are grounded.

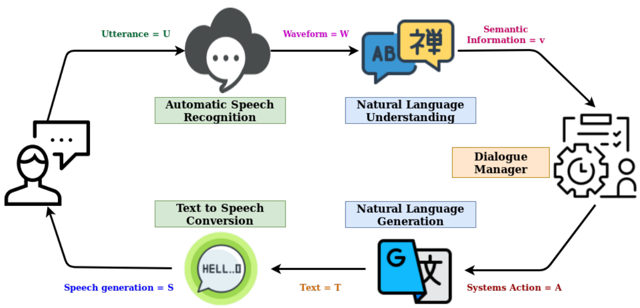

Figure 2: A traditional Spoken Dialogue Management System.

2.2 Mediation of Joint Intention

The agent and the user have to use a medium to communicate their joint intention, and to negotiate goals g and commitments O. In order to do that, we formalise a finite-state negotiation dialogue strategy with following dialogue acts: *(offer, counter-offer, accept, and reject)*. The dialogue strategies will be implemented with a spoken dialogue management system (SDS) 2.

Traditionally, an SDS consists of speech synthesis that recognises and generates speech, natural language understanding (NLU) transforms the human generated natural language to knowledge for the machine. A dialogue manager (DM) makes the decision based on the NLU and other components such as previous history, database etc, and natural language generation (NLG) receives the decision from the DM, transforms it to human understandable format and sends to speech synthesizer.

When the user generates its first utterance, it is transformed from speech to text and arrives at the NLU component. The NLU transforms the text to knowledge (semantic roles) and assigns a dialogue act *offer* [2]. An offer from the user instantiates the plan π by performing plan recognition, and creates a joint intention as described by $JointIntend$.

We define five dialogue acts $Offer_a$, $Offer_g$, *Counter-offer, Accept* and *Reject* with which both user and DM can mediate the intention's goal g and commitments O. Table 1 contains the effects of these dialogue acts with respect to three sets: θ is a set of offers, R and O are respectively the sets of rejected and accepted commitments. An $Offer_a$ represents offer about an action, $Offer_g$ indicates offer about a goal, *Counter-offer* is an action a_1 is not accepted and an alternative a_2 is instead proposed. An *Accept* and *Reject* can be used to accept or reject proposed commitments.

[2] At this stage we define a finite state SDS and only the user can initiate a dialogue only using *offer* together with a goal or an action.

Dialogue Act	Precondition	Effect
$Offer_a, x, a$	$a \notin \theta \land a \notin R \land a \notin R$	$a \in \theta$
$Offer_g, x, \hat{g}$	$\emptyset$	$g = \hat{g}$
$Counter\text{-}offer, x, a_1, a_2$	$a_1 \in \theta\ a_2 \notin R \land a_2 \notin O$	$a_1 \notin \theta \land a_2 \in \theta$
$Accept, x, a$	$a \in \theta$	$a \notin \theta \land a \in O$
$Reject, x, a$	$a \in \theta$	$a \notin \theta \land a \in R$

Table 1: Speech acts for the SDS that allow to mediate actions with respect to the sets of offered, rejected and accepted commitments.

Since a dialogue policy based on Finite-State-Machines is not realizable as it would need to consider all possible intentions, also based on the state of the task, we propose to learn the DM dialogue policy with Reinforcement Learning methods. This approach is not new in the context of dialogue management, and by this method the user is simulated by an Agenda [12].

2.3 Learning the agent strategy with Reinforcement Learning (RL)

At every turn, a Q-Network [18] evaluates the current inferred π together with the actions in the sets θ, R and O and the current PDDL state, selecting which dialogue act to perform by an ϵ-greedy policy computed on the dialogue acts expected return. 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:

$$R = -\alpha T + \beta \frac{\overline{\pi} \cap \pi}{\overline{\pi} \cup \pi} + \gamma \frac{C(\overline{\pi}, \overline{g})}{C(\pi, \overline{g})} \tag{5}$$

where $\overline{\pi}$ and $\overline{g}$ form the user's original desired joint intention (held by the Agenda). 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 resulting 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. Thus, the the DM learns to mediate and improve the unobservable user intention $\overline{\pi}, \overline{g}$.

3 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 components together with a simple user Agenda. The Agenda is based on PDDL and simulates how the user would modify the joint intention during its turn, while having as objective a randomly generated joint intention.

Initial experiments gave positive results, in the sense that the RL 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: 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 with SDS is still to be implemented. For early prototypes of the system we plan to implement the dialogue manager as described in Section 2.2. Later versions could see the implementation of a more complete SDS through for example a POMDP model [9]. 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. Finally, investigation about the soundness of this approach in real scenarios for example in user studies is still to be performed.

ACKNOWLEDGMENTS

This work has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 721619 for the SOCRATES project.

REFERENCES

[1] Philip R Cohen. 2019. Foundations of Collaborative Task-Oriented Dialogue: What's in a Slot?. In *Proceedings of the 20th Annual SIGdial Meeting on Discourse and Dialogue*. 198–209.

[2] Philip R Cohen and Hector J Levesque. 1990. Intention is choice with commitment. *Artificial intelligence* 42, 2-3 (1990), 213–261.

[3] Sandra Devin and Rachid Alami. 2016. An implemented theory of mind to improve human-robot shared plans execution. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 319–326.

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

[5] Barbara Grosz and Sarit Kraus. 1996. Collaborative plans for complex group action. *Artificial Intelligence* (1996).

[6] Bunt Harry, Petukhova Volha, Traum David, and Alexandersson Jan. 2017. *Dialogue Act Annotation with the ISO 24617-2 Standard*. Springer International Publishing, 109–135.

[7] Ece Kamar, Ya'akov Gal, and Barbara J Grosz. 2009. Incorporating helpful behavior into collaborative planning. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. Springer Verlag.

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

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

[10] 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.

[11] Timothy W. Rauenbusch and Barbara J. Grosz. 2003. A Decision Making Procedure for Collaborative Planning. In *Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS '03)*. Association for Computing Machinery, 1106–1107.

[12] Jost Schatzmann, Blaise Thomson, and Steve Young. 2007. Statistical user simulation with a hidden agenda. In *Proceedings of the 8th SIGdial Workshop on Discourse and Dialogue*. 273–282.

[13] 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.

[14] Ira A Smith, Philip R Cohen, Jeffrey M Bradshaw, Mark Greaves, and Heather Holmback. 1998. Designing conversation policies using joint intention theory. In *Proceedings International Conference on Multi Agent Systems (Cat. No. 98EX160)*. IEEE, 269–276.

[15] 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.

[16] 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.

[17] Raimo Tuomela. 2005. We-intentions revisited. *Philosophical Studies* 125, 3 (2005), 327–369.

[18] 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.