Intelligent Decision Support in Human Team Planning

Joseph Kim

Computer Science and Artificial Intelligence Laboratory
Massachusetts Institute of Technology
joseph_kim@csail.mit.edu

1 Introduction

Human team planning is an integral component in almost all collaborative and organized work environments. Human team planning sessions (i.e. meetings), however, often suffer from inefficiencies such as getting off topic, losing track of goals, inadequate preparation, and a lack of shared understanding between the team members. An estimated amount of $54 million to $3.4 billion is lost annually due to general meeting inefficiencies [Romano and Nunamaker Jr., 2001].

In addition to the *social* difficulties of team planning, the underlying planning task itself could be challenging for humans to solve. In resource optimization and mission planning, humans tend to employ greedy heuristics, overlook constraints, and lose situational awareness when working with a large amount of data [Miller, 2004; Maule, 2010]. In time and safety-critical domains like disaster response, the complexities of both the planning task and social factors are heightened where mishaps in planning can trigger high costs during plan execution.

In order to alleviate the human challenges of team planning, researchers in AI planning have been developing human-agent planning systems (also called human-in-the-loop planning or intelligent decision support (IDS)), where the goal is to reduce cognitive load and help human teammates produce high-quality plans prior to execution. However, there are two primary ongoing challenges. First, much focus in human-agent planning has been limited to dyadic interaction between a single human and a single IDS agent. Humans, however, often plan in groups (team planning makes up more than 70% of our time at work [Romano and Nunamaker Jr., 2001]), requiring existing IDS systems to incorporate social science models of group decision-making and teamwork. The second problem is that a large body of existing IDS systems have been designed to be domain-specific and/or coupled to a specific user interface. Generalizable IDS models, designed to monitor human team planning sessions, and evaluations of such models across various planning domains require much attention.

The aim of my thesis is to develop computational models for an IDS agent participating in human team planning sessions. I intend to develop novel techniques for an IDS agent that can 1) predict occurrences when the team's level of shared understanding is weak, 2) infer the team's intended plan directly from their planning conversations, and 3) sug-
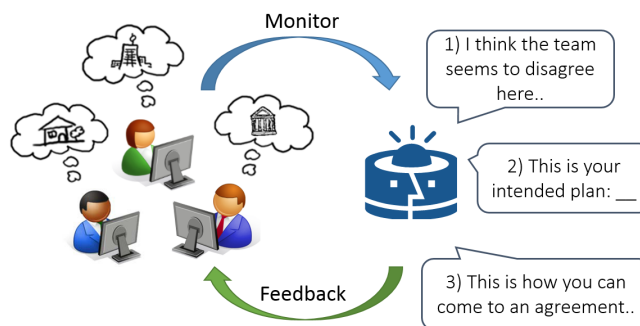

Figure 1: A visual illustration of an agent monitoring and providing feedback in human team planning. The three IDS tasks investigated in my thesis are expressed by the agent.

gest a resolution strategy when plan differences occur among human teammates. Figure 1 provides an illustration of such IDS functionalities that would significantly improve our ability to coordinate and generate high-quality plans, not only in disaster response scenarios, but in everyday work meetings as well.

2 Improving Team's Shared Understanding

Upon concluding a meeting, team members can occasionally leave with different understandings of what had been discussed. Detecting inconsistencies in understandings is a desired capability for an IDS agent that can eventually provide feedback to spur stronger team shared understanding. I developed a computational model to automatically predict the team's consistency of understanding in meetings. The model incorporated a multi-modal feature set of textual dialogue acts and visual head gestures. I utilized a feature set inspired by social science studies in capturing the dynamics of the team agreement process [Di Eugenio *et al.*, 2000].

By using one of the largest publicly available meeting datasets and a machine learning algorithm, my model learned discriminative sequences of features leading to either strong or weak levels of shared understanding [Kim and Shah, 2014]. I then implemented the learned model within an IDS agent that participated in live, web-based meetings (communication via text chat). At the end of the meeting, the agent suggested review of discussions predicted to have weak lev-

els of shared understanding. Through human subject experiments, I evaluated the utility of such intelligent review system and observed a statistically significant increase of 18% in the objective measures of team's shared understanding [Kim and Shah, 2016].

3 Inferring the Team's Intended Plan

The IDS agent, however, must do more than predict whether or not the human team is in agreement. In order to provide a detailed level of plan evaluation and critique, the agent needs to maintain a formal representation of the underlying *plan* in which the team is in agreement to. In my thesis, I develop a novel plan inference model to infer the team's final plan (represented using the PDDL language) directly from a processed form of their planning conversation. Reasoning from this natural form of communication enables the model to generalize across various planning domains and remove the dependency for any manual plan transcribing from humans.

The key to my approach involves leveraging features from the team's dialogue pattern and using planning domain models to efficiently search the large space of possible plans. I employ reinforcement learning, where the agent learns to predict final plans without the use of traditional, supervised dataset (e.g. post-meeting summaries, plan execution traces). While only using planning feedback as its source of supervision, my approach achieves performance on par with that of a supervised baseline. I empirically demonstrate that the learned model can be applied to successfully infer team plans within a novel planning domain, and that performance is less sensitive to incomplete model specifications than prior art. I will present the work at the IJCAI-18 main technical track [Kim *et al.*, 2018].

There are several extensions I seek to explore in future work. It is important to consider the effect of social roles and more complicated social interactions during human team planning. Different types of expertise, authority and group sizes represent possible additional features to include and analyze in the plan inference model (e.g. an option expressed by an advisor, compared to that of an advisee, may be more likely to be included in the final plan). Second, I seek to develop an *active* plan inference agent designed to query information from human teammates to better infer and support their plans. For such an interaction, a socially-aware agent may be developed, one that understands when it should behave proactively or passively, acting in response to the team's evolving social state (e.g. level of team agreement, stress, or confusion).

4 Explaining Multiple Plan Differences

When team members' proposed plans differ (either due to inconsistency in their shared understanding or due to deliberate disagreement), it would be helpful to the team's resolution process if the IDS agent can intuitively explain *how* the plans are different. In my thesis, I investigate the use of linear temporal logic (LTL) as an interpretable language for providing explanations on plan differences. In my prior work, I conducted a human subject experiment and found that a large portion of human high-level strategies can be converted into

LTL formulas in which the low-level planner can use to generate human-interpretable plans [Kim *et al.*, 2017]. Inversely, LTL formulas can be translated back into high-level natural language descriptions. Based on such findings, I will develop a learning model that given an arbitrary pair of divergent plans, intelligently selects a sequence of LTL formulas to maximize human interpretability of understanding plan differences. I will compare and analyze the differences between an explanation model learned from human labels, against an unsupervised model learned from planning metrics alone.

5 Conclusion

We are witnessing an increasing amount of online collaboration tools and team messaging software such as Slack. There is a growing level of interest in developing cognitive AI assistants with such platforms. IDS systems for human team planning will be highly applicable in both emergency response and everyday work environments, and in future may exist as an embodied agent participating in physical meetings. Attending IJCAI, the premier conference in AI, will help me to connect with top researchers in the broad field of human-agent planning, where I hope to gain invaluable insights and collaborations to realize the full potential of my work.

References

[Di Eugenio *et al.*, 2000] Barbara Di Eugenio, Pamela W Jordan, Richmond H Thomason, and Johanna D Moore. The agreement process: An empirical investigation of human–human computer-mediated collaborative dialogs. *Int. Jour. Human-Computer Studies*, 53(6):1017–1076, 2000.

[Kim and Shah, 2014] Joseph Kim and Julie A Shah. Automatic prediction of consistency among team members' understanding of group decisions in meetings. In *Int. Conf. Systems, Man and Cybernetics*, 2014.

[Kim and Shah, 2016] Joseph Kim and Julie A Shah. Improving team's consistency of understanding in meetings. *Trans. Human-Machine Systems*, 46(5):625–637, 2016.

[Kim *et al.*, 2017] Joseph Kim, Christopher J Banks, and Julie A Shah. Collaborative planning with encoding of users' high-level strategies. In *AAAI*, pages 955–962, 2017.

[Kim *et al.*, 2018] Joseph Kim, Matthew E Woicik, Matthew C Gombolay, Sung-Hyun Son, and Julie A Shah. Learning to infer final plans in human team planning. In *IJCAI - To Appear*, 2018.

[Maule, 2010] A John Maule. Can computers help overcome limitations in human decision making? *Intl. Jour. of Human–Computer Interaction*, 26(2-3):108–119, 2010.

[Miller, 2004] C. Miller. Modeling human workload limitations on multiple UAV control. In *Human Factors*, pages 526–527, 2004.

[Romano and Nunamaker Jr., 2001] N.C. Romano and J.F. Nunamaker Jr. Meeting analysis: findings from research and practice. In *Int. Conf. System Sciences*, 2001.