Agent-human Coordination with Communication Costs under Uncertainty*

(Extended Abstract)

Asaf Frieder[1], Raz Lin[1] and Sarit Kraus[1,2]
[1] Department of Computer Science
Bar-Ilan University
Ramat-Gan, Israel 52900
{linraz,sarit}@cs.biu.ac.il
[2] Institute for Advanced Computer Studies
University of Maryland
College Park, MD 20742 USA

ABSTRACT

As agents' technology becomes increasing more prevalent, coordination in mixed agent-human environments becomes a key issue. Agent-human coordination is becoming even more important in real life situations, where uncertainty and incomplete information exists and communication is costly. While abundant research has focused on aspects of computerized teamwork, little attention has been given to the issues raised in teams that consist of both computerized agents and people. In this paper we focus on teamwork between an agent and a human counterpart and present a novel agent designed to interact proficiently with people. In extensive simulations we matched our agent with people and compared it with another state-of-the-art agent. Our results demonstrate the significant improvement in coordination when our agent is involved.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence

General Terms

Experimentation

Keywords

human-robot/agent interaction, POMDPS, uncertainty, teamwork

1. INTRODUCTION

More and more agents are deployed in mixed agent-human environments and are expected to interact efficiently with people. Such settings may include uncertainty and incomplete information. Communication, which can be costly, might be available for the parties to assist in obtaining more information in order to build a good model of the world. Efficient coordination in teams between agents and people is the key component for turning their interaction into a successful one. The importance of coordination between agents and people only increases in real life situations, in which uncertainty and incomplete information exist.

Teamwork has been the focus of abundant research in the multi-agent community. However, while research has focused on decision theoretic framework, communication strategies and multi-agent policies (e.g., [2]), only some focus has been on the issues raised when people are involved as part of the team [3]. Our work focuses on efficient coordination between agents and people with communication costs and uncertainty. We model the problem using DEC-POMDPs (Decentralized Partially Observable Markov Decision Process) [1]. The problem involves a team having a joint reward (goals), while each team member has only partial observations of the state of the world. Thus, even if information exists, it only provides partial support as to the state of the world, making it difficult to construct a reliable view of the world without coordinating with other teammates. To validate the efficacy of our agent, we chose the Colorado/Wyoming domain, which was first introduced by Roth *et al.* [2] and offered as a benchmark for evaluation of communication heuristics in multi-agent POMDPs.

While there are studies that focus on DEC-POMDPs, most of them pursue the theoretical aspects of the multi-agents aspect, and do not deal with the fact that people can be part of the team [2]. Zuckerman *et al.* [4] improved coordination with humans using focal points. We, however, focus on the problem of improving coordination between an agent and people by means of shared observations. The addition of communication only increases the challenge, making the adaptation of their model far from straightforward. Our novelty also lies in introducing an agent capable of successfully interacting with a human counterpart in such settings. The agent is adaptable to the environment and people's behavior, and is able to sophisticatedly decide which information to communicate to the other team member based on the communication cost and the possible effects of this information on its counterpart's behavior.

2. COORDINATION WITH COMMUNICATION COSTS IN DEC-POMDPS

A DEC-POMDP [1] model separates the resolution of the problem into time steps in which the agents choose actions simultaneously. These actions can have deterministic or non-deterministic effects on the state. Following these actions, each team member receives an additional observation of the world state. The state transition and the joint reward function are dependent on the joint actions of all agents.

*This research is based upon work supported in part by the U.S. Army Research Lab and Research Office grant number W911NF-08-1-0144, ARO grants W911NF-09-1-0206 and W911-NF-11-1-0344 and by ERC grant #267523.

Appears in: *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012)*, Conitzer, Winikoff, Padgham, and van der Hoek (eds.), 4-8 June 2012, Valencia, Spain.

We focus on POMDPs in which the team consists of two agents and the team members are able to communicate with each other (e.g., [2]). As communication is costly we limit the communication messages to include only self observations. This can also be supported in real settings where limitations occur to prevent lengthy communications that can breach the integrity of the team members (e.g., surrendering their locations). By sharing their observations, the team members can avoid uncoordinated actions caused by contradictory private knowledge, allowing them to build a coherent and a concise view of the world states faster.

A naïve approach for team communication is sharing all information among themselves. Once all the information is shared, finding the optimal joint action becomes a simple POMDP problem that each team member can solve in parallel. Then, each member can perform the action assigned to it in the joint action plan, described by the POMDP policy, for their joint belief. However this solution is only optimal if two assumptions hold. First, that there is no cost associated with communication. Second, that all team members consider the same joint actions to be optimal (by using the same POMDP policy). As this is hardly the case in real settings, existing agents might fail when matched with people. Our agent's design takes these considerations into account to achieve proficient interaction with people.

3. AGENT DESIGN

When coordinating with someone else, it is hard to predict with full certainty what the other team member (especially if it is a human partner) will do. The task is even harder if the agent interacts with someone only once and not repeatedly. Thus, an efficient agent working with people needs, amongst other things, to approximate what percentage of the population will perform each action based on the existing partial observations. Our agent interacts with the same counterpart only once and thus its design tries to tackle the challenge by generating a good model of the population based on 445 people who played the game. Our agent uses a neural network which outputs the probabilities of the other team member taking an action based on features that encod the agent beliefs, past actions and communication- and position-related information. We coin our agent *TMDC* (standing for team modeling with decentralized communication).

4. DESIGNING THE AGENT'S STRATEGY

The general design of the agent's strategy consists of building a POMDP using the prediction of the human behavior described beforehand. This is done as when interacting with people we cannot ensure mutual predictability. Thus, *TMDC* uses its model, and not the shared belief, to predict what will be its counterpart's behavior. In addition, *TMDC* chooses its action based on all its knowledge (which also includes private knowledge), and only communicates in order to influence the actions of the other teammate. Given all previously shared observations, the agent evaluates an action by considering all possible results, calculating immediate rewards and using offline estimation of future rewards. This evaluation is then used by a hill climbing heuristic that finds which observations (taken from the set of all observations, including shared observations) can maximize the score of the team and hence should be shared. We continue to describe the agent's strategy in detail.

5. EXPERIMENTS

The experiments were conducted on the Colorado/Wyoming domain and were conducted using the Amazon Mechanical Turk service (AMT). This framework allows publishing of tasks designated for people all around the world. We prohibited multiple participation by the same people. The players were provided with a manual of the game before their participation. Although the manual is very detailed, we took great care not to give strategic advice. We then required that each worker pass a short multiple choice test to verify that they read the manual and understood the game. The player received a bonus based on the score of the team, if it was positive. We ensured that the costs and penalties of the game would have a meaningful effect on the player even if the team did not gain the reward for a successful signal.

We matched 64 people with our *TMDC* agent, with a state-of-the-art agent *PDCS* [2] and with 64 other people (*PDCS* was designed to coordinate well with multi-agent teams). The results demonstrate that our agent significantly outperforms the *PDCS* agent ($p < 0.001$) when matched with people (52.84 as compared to 17.5). Interesting also that the human-human team achieved a score of only 27.18. While this score generated no significance difference compared to the *PDCS* results, the *TMDC*-human teams achieved significantly higher scores ($p < 0.003$) than the human-human teams.

6. CONCLUSIONS

Settings in which hybrid teams of people and automated agents need to achieve a common goal are becoming more common in today's reality. Communication in such situations is a key issue for coordinating actions. As communications is costly and sometimes even limited (e.g., due to security issues or range limitations) it becomes of great essence to devise an efficient strategy to utilize communication. This paper presented a novel agent design that can proficiently coordinate with people under uncertainty while taking into account the cost of communication.

Our agent was specifically designed taking into account the fact that it interacts with people and was also evaluated with people. Experiments with more than 300 people demonstrated how it outperforms the state-of-the-art agent. One of the main factors accounting for the success of our agent is the understanding that it requires a good model of the counterpart to generate an efficient strategy.

This paper is only part of a new and exciting journey. Future work warrants careful investigation on improving the prediction model of people's behavior. In addition we will investigate settings in which even more limited information is available to the team members. In such situations the challenge is on the understanding of the abstract model that is available and how to utilize communication's strategies for efficient coordination that will allow increasing the accuracy of the model.

7. REFERENCES

[1] D. Bernstein, R. Givan, N. Immerman, and S. Zilberstein. The complexity of decentralized control of markov decision processes. *Mathematics of Operations Research*, 27(4):819–840, 2002.

[2] M. Roth, R. Simmons, and M. Veloso. What to communicate? execution-time decision in multi-agent POMDPs. *Distributed Autonomous Robotic Systems*, 2006.

[3] A. van Wissen, Y. Gal, B. Kamphorst, and V. Dignum. Human–agent teamwork in dynamic environments. *Computers in Human Behavior*, 28:23–33, 2012.

[4] I. Zuckerman, S. Kraus, and J. S. Rosenschein. Using focal points learning to improve human-machine tactic coordination. *JAAMAS*, 22(2):289–316, 2011.