



DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

To Share or not to Share? The Single Agent in a Team Decision Problem

The Harvard community has made this article openly available. [Please share](#) how this access benefits you. Your story matters.

Citation	Amir, Ofra, Barbara Grosz, and Roni Stern. "To Share or not to Share? The Single Agent in a Team Decision Problem." Presented at the 28th AAAI Conference on Artificial Intelligence, Québec City, Québec, Canada, July 27-31, 2014.
Published Version	http://www.aaai.org/ocs/index.php/AAAI/AAAI14/paper/viewFile/8433/8760
Accessed	May 22, 2018 11:18:13 PM EDT
Citable Link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:14169434
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP

(Article begins on next page)

To Share or Not to Share? The Single Agent in a Team Decision Problem

Ofra Amir and Barbara J. Grosz

School of Engineering and Applied Sciences
Harvard University
oamir@seas.harvard.edu, grosz@seas.harvard.edu

Roni Stern

Department of Information Systems Engineering
Ben-Gurion University of the Negev
roni.stern@gmail.com

Abstract

This paper defines the “Single Agent in a Team Decision” (*SATD*) problem. *SATD* differs from prior multi-agent communication problems in the assumptions it makes about teammates’ knowledge of each other’s plans and possible observations. The paper proposes a novel integrated logical-decision-theoretic approach to solving *SATD* problems, called *MDP-PRT*. Evaluation of *MDP-PRT* shows that it outperforms a previously proposed communication mechanism that did not consider the timing of communication and compares favorably with a coordinated *Dec-POMDP* solution that uses knowledge about all possible observations.

Introduction

This paper defines a new multi-agent decision problem, the “Single Agent in a Team Decision” (*SATD*) problem, which may be described informally as follows: An individual collaborating in a multi-agent team obtains new information, unanticipated at planning time. This (single) agent has incomplete knowledge of others’ plans. It must decide whether to communicate this new information to its teammates, and if so, to whom, and at what time. *SATD* differs from previously studied multi-agent communications problems in that it does not assume complete knowledge of other agents’ plans or policies nor that all observations are knowable in advance. It assumes instead that agents have some knowledge of each other’s intentions and plans which can be used to reason about information sharing decisions.

We are investigating *SATD* in the context of developing computer agents to support care teams for children with complex conditions (Amir et al. 2013). Care teams for children with complex conditions involve many providers – a primary care provider, specialists, therapists, and non-medical care givers. The care team defines a high-level care plan that describes the main care goals, but there is no centralized planning mechanism that generates a complete plan for the team or that can ensure coordination. Caregivers are unaware of their collaborators’ complete plans, yet their individual plans often interact. Communicating relevant information among team members is crucial for care to be coordinated and effective, but doing so is costly and often insufficient in practice.

Copyright © 2014, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

Developing agents that are capable of supporting information sharing in such teams is beyond the current state-of-the-art in multi-agent systems. *BDI* approaches to multi-agent planning, e.g. *STEAM* (Tambe 1997), often base their communication mechanisms on theories of teamwork and collaboration (Grosz and Kraus 1996; Cohen and Levesque 1990). These approaches, however, typically do not reason about uncertainty and utilities that are prevalent in the healthcare domain. Decision-theoretic approaches to multi-agent communication typically reason about communication within a *POMDP* framework, e.g., (Goldman and Zilberstein 2003; Roth, Simmons, and Veloso 2006; Pynadath and Tambe 2002). However, these approaches assume that all possible observations are known in advance and that the team has a joint policy. In contrast, in the healthcare domain new information that was unexpected at planning time is often observed. In addition, care providers only agree on high-level goals and team members individually plan ways to accomplish the tasks for which they are responsible. Therefore, it cannot be assumed that agents know the complete plans or policies of their teammates.

To address *SATD*, we propose a novel, integrated Belief-Desire-Intention (*BDI*) and decision-theoretic (*DT*) representation, called “*MDP-PRT*” that builds on the strengths of each approach. Evaluation of an agent using *MDP-PRT* shows that it outperforms the inform algorithm proposed by Kamar et al. (2009). In addition, we compared the agent’s performance with that of *Dec-POMDP* policy that was informed about all possible observations and show it obtains close results to those obtained by this *Dec-POMDP* policy despite lacking a coordinated policy that considers all possible observations.

The paper makes three contributions: (1) it formally defines the *SATD* communication problem; (2) it proposes a new representation (*MDP-PRT*) that enables agents to reason about and solve the *SATD* communication problem, and (3) it demonstrates the usefulness of this representation.

Problem Definition

In this section, we formally define the *SATD* problem. *SATD* arises in the context of a group activity of a team of agents. We assume the group’s plan meets the SharedPlans specification for collaborative action (Grosz and Kraus 1996). Two particular properties of SharedPlans are impor-

tant: plans, which decompose into constituent tasks, may be only partially specified and agents do not know the details of constituent tasks for which they are not responsible. An instance of the SATD problem is represented by a tuple $\langle a_i, A_{-i}, b_{SP}, V, o^*, \varphi_{comm}, C \rangle$.

a_i is the agent that observes new information. o^* is the new information a_i obtained. A_{-i} are the other agents that are part of the team. b_{SP} is a_i 's beliefs about the SharedPlan of the team (a_i knows its own plans but he is uncertain about others' plans). V is the utility function; its value is the utility of completed constituent tasks. We assume a fully cooperative setting and all agents share the same utility. φ_{comm} is a function that produces a modified b'_{SP} under the assumption that the agents in A_{-i} (or a subset of them) are informed about o^* . C is the cost of communicating o^* . SATD is the problem of a_i determining whether to communicate o^* to agents in A_{-i} and if so, at what time.

Approach

This section proposes a solution to the 2-agent SATD problem $\langle a_1, a_2, b_{SP}, V, o^*, \varphi_{comm}, C \rangle$, where an agent a_1 learns new information o^* and needs to reason whether and when to communicate o^* to a_2 . To solve the 2-agent SATD communication problem, we use an MDP in which the states explicitly represent a_1 's beliefs about a_2 's plans. The choice of representation for the agent's beliefs b_{SP} is key, as it affects the way b_{SP} can be revised and therefore the computational efficiency of solving the MDP. Our approach uses a Probabilistic Receipt Tree (PRT) (Kamar, Gal, and Grosz 2009) to represent b_{SP} .

In the MDP-PRT $\langle A, S, R, Tr, s_0 \rangle$, A includes two actions *inform* (communicating o^*) and *-inform* (not communicating o^*). Each state in S encompasses a_1 's beliefs about a_2 's plans (i.e., the PRT corresponding to b_{SP}). We denote a state by b_{SP} . The initial state b_{SP_0} corresponds to a_1 's initial beliefs about the SharedPlan. The reward function is a function of V and C : the reward for a state b_{SP} is the value of the constituent tasks completed in the last time step minus the cost of communication if a_1 chose to inform a_2 . The transition function, $Tr(b_{SP}', a, b_{SP})$, defines the probability of reaching state b_{SP}' , when taking action a in state b_{SP} . a_1 's belief may change for two different reasons. First, if a_1 communicates o^* to a_2 , then b_{SP} changes to $\varphi_{comm}(b_{SP}, o^*)$. Second, as a_2 executes actions in its constituent plans, a_1 may "observe" a_2 's actions or results of those actions and learn more about a_2 's plans. To reflect this reasoning, we define an additional function, $\varphi_{obs}(b_{SP})$. This function takes as input b_{SP} and returns the set of next expected beliefs b_{SP}^{next} and their probabilities $\Pr(b_{SP}^{next})$.

An optimal single agent communication policy can be computed using any MDP solver, e.g. value iteration algorithm used in our implementation.

Results

We tested the MDP-PRT agent using a modified version of the Colored Trails (CT) game used by Kamar et al. (2009). The first experiment compared the performance of the MDP-PRT agent with that of a PRT agent using the inform algo-

rithm of Kamar et al. (2009). Table 1 shows the average utility achieved by the agents in experiments with different communication costs. The MDP-PRT agent outperforms the PRT agent across all configurations and communication costs. This is because the MDP-PRT considers the possibility of communicating at a later time, while the PRT inform algorithm is myopic.

Comm. Cost	5	10	20	30
PRT	67.67	65.17	60.17	55.17
MDP-PRT	77	75.57	72.83	70.03

Table 1: Average utility for PRT Inform and MDP-PRT.

Key aspects of SATD are that agents cannot anticipate a_1 observing o^* nor form a coordinated policy. This limits the possible utility achievable in SATD, as a coordinated policy that takes into account knowledge of possible observations during planning time can lead to better performance. To evaluate the loss in utility as a result of this missing knowledge, we compare the MDP-PRT with a joint policy generated using a Dec-POMDP that is knowledgeable of the possible observations at planning time. As expected, the Dec-POMDP always performs better as a result of its additional information at planning time (see table 2). However, the average utility achieved by MDP-PRT is still within 15% of the optimal utility.

Comm. Cost	5	10	25	50
Dec-POMDP	101.8	101.46	100.2	98.13
MDP-PRT	100.52	99.08	94.38	87.5

Table 2: Average utility for Dec-POMDP and MDP-PRT.

Acknowledgements

The research reported in this paper was supported in part by a grant from the Nuance Foundation.

References

- Amir, O.; Grosz, B. J.; Law, E.; and Stern, R. 2013. Collaborative health care plan support. In *Proceedings of the 12th international conference on Autonomous agents and multi-agent systems*, 793–796.
- Cohen, P., and Levesque, H. 1990. Intention is choice with commitment. *Artificial intelligence* 42(2):213–261.
- Goldman, C. V., and Zilberstein, S. 2003. Optimizing information exchange in cooperative multi-agent systems. In *Proceedings of the second international joint conference on Autonomous agents and multiagent systems*, 137–144.
- Grosz, B., and Kraus, S. 1996. Collaborative plans for complex group action. *Artificial Intelligence* 86(2):269–357.
- Kamar, E.; Gal, Y.; and Grosz, B. 2009. Incorporating helpful behavior into collaborative planning. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems*, 875–882.
- Pynadath, D. V., and Tambe, M. 2002. The communicative multiagent team decision problem: analyzing teamwork theories and models. *Journal of Artificial Intelligence Research* 16(1):389–423.
- Roth, M.; Simmons, R.; and Veloso, M. 2006. What to communicate? execution-time decision in multi-agent POMDPs. *Distributed Autonomous Robotic Systems* 7 177–186.
- Tambe, M. 1997. Agent architectures for flexible practical teamwork. *AAAI* 97(1):997.