Team Formation and Communication Restrictions in Collectives

Adrian K. Agogino
The University of Texas
ECE Department
Austin, TX 78712
agogino@ece.utexas.edu

Kagan Tumer
NASA Ames Research Center
Mailstop 269-3
Moffett Field, CA 94035
kagan@ptolemy.arc.nasa.gov

ABSTRACT

A collective of agents often needs to maximize a "world utility" function which rates the performance of an entire system, while subject to communication restrictions among the agents. Such communication restrictions make it difficult for agents which try to pursue their own "private" utilities to take actions that also help optimize the world utility. Team formation presents a solution to this problem, where by joining other agents, an agent can significantly increase its knowledge about the environment and improve its chances of both optimizing its own utility and that its doing so will contribute to the world utility. In this article we show how utilities that have been previously shown to be effective in collectives can be modified to be more effective in domains with moderate communication restrictions resulting in performance improvements of up to 75%. Additionally we show that even severe communication constraints can be overcome by forming teams where each agent of a team shares the same utility, increasing performance an additional 25%. We show that utilities and team sizes can be manipulated to form the best compromise between how "aligned" an agent's utility is with the world utility and how easily an agent can learn that utility.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Multiagent Systems

General Terms

Performance, Economics, Experimentation, Theory

Keywords

Reinforcement learning, MAS, Q-learning

1. INTRODUCTION

Many methods exist for coordinating the actions of a multiagent system when the agents can fully communicate with one another [3, 4]. However, many problems impose communication restrictions among the agents, rendering the coordination problem more difficult [1]. Examples of these problems, include controlling collections of rovers, constellations of satellites and packet routers, where an agent may only be able to directly communicate with a small number of other agents. In all of these problems, the collective's designer faces the following difficult task:

- ensuring that, as far as the provided "world utility function" is concerned, the agents do not work at cross-purposes (i.e., making sure that the private utilities of the agents and the world utility are "aligned").
- ensuring that agents can achieve their private utilities when they do not have access to a broad communication network giving them access to global information.

These tasks can be addressed with the theory of collectives which has been successfully applied to multiple domains including packet routing over a data network, the congestion game known as Arthur's El Farol Bar problem [4], and the coordination of multi-rovers in learning sequences of actions.

The thoery of collectives is concerned with the **world utility** $G(z)$, which is a function of the full worldline, z. The problem at hand is to find the z that maximizes $G(z)$. In addition to G, for each agent η, there is a **private utility function** g_η. The agents act to improve their individual private functions, even though, we, as system designers are only concerned with the value of the world utility G. An important property we want a private utility to have is **factoredness** with respect to G, intuitively meaning that an action taken by an agent that improves its private utility also improves the world utility. In addition to being factored we want the agents' private utility functions to have high **learnability**, intuitively meaning that an agent's utility should be sensitive to its own actions and insensitive to actions of others. As a trivial example, any "team game" in which all the private functions equal G is factored, but has low learnability since all the agents' actions have a significant effect on the value of G.

Consider **difference** utilities, which are of the form:

$$DU\eta \equiv G(z) - G(CL_\eta(z)) \tag{1}$$

where $CL_\eta(z) = (z_{\hat{\eta}}, \ell_\eta)$ is a pre-fixed **clamping parameter** ℓ_η chosen from among η's legal or illegal moves. Such difference utilities are factored no matter what the choice of clamping parameter because the second term does not depend on η's state [4]. Furthermore, they usually have far better learnability than does a team game because the second term of DU which removes a lot of the effect of other agents (i.e., noise) from η's utility.

AAMAS'03, July 14–18, 2003, Melbourne, Australia.
ACM 1-58113-683-8/03/0007.

1.1 Communication Restrictions and Teams

Mathematically we will represent the communication restrictions as elements of the worldline that are not observable. Given a worldline z, we can decompose it into an observable components, z^o, and hidden components, z^h (we will denote the concatenated state z by $z = (z^o, z^h)$). If the DU depends on any component of z^h then we cannot compute it directly. Instead there are several approximations to the DU that vary in their balance between learnability and factoredness. In this paper we propose 4 approximations [1]:

$$\begin{aligned} BTU_\eta(z) &= G(z) - G(CL_\eta(z^o, \vec{0})) && (2)\\ TTU_\eta(z) &= G((z^o, \vec{0})) - G(CL_\eta(z^o, \vec{0})) && (3)\\ BEU_\eta(z) &= G(z) - G(CL_\eta(z^o, E[z^h|z^o])) && (4)\\ EEU_\eta(z) &= G((z^o, E[z^h|z^o])) - \\ & \quad\; G(CL_\eta(z^o, E[z^h|z^o])) && (5) \end{aligned}$$

where $\vec{0}$ is the vector whose components are all zero, CL_η clamps all components of agent η to the zero vector, and $E[]$ is the expectation operator. Note that the BTU and BEU assume that the true world utility can be produced despite the communication restriction. These two utilities are also factored since they are in the form of equation 1, however they may not be very learnable since the second term uses different information from the first, causing less noise to be subtracted out. EEU does not have this problems, and with a good estimate of z^h it may still be close to being factored.

As discussed above, communication restrictions can have serious negative effects on the utility functions of the agents. One way to remedy this situation is to let agents form "teams" which "share" information [2]. In this paper a team is defined as an aggregation of agents where each agent: (1) belongs to one and only one team, (2) receives the utility of the team, and (3) shares information with its team members.

2. EXPERIMENTAL RESULTS

We conducted a series of experiments on a generalized version of the El Farol Bar Problem described in [4]. The first set of experiments were conducted without teams (team size = 1). Figure 1 shows the performance of the four utilities with different levels of communication. With high communication levels, all the utilities converge to the DU. When communication is very low, the BTU and BEU have the best performance because their first term G is not affected by the communication restriction, and converge to G when communication is zero. However these utilities have trouble incorporating additional knowledge and cannot do better than G when performance below the 50% communication level. At most communication levels, the EEU performs the best, since it is the most learnable and is very close to being factored. TTU performs the worst at most communication levels since it is not close to being factored.

Even using the best utility, EEU, a high level of performance cannot be achieved if the communication level is too low. However if agents can form small teams where information sharing is allowed between team members, good performance is possible even when communication between

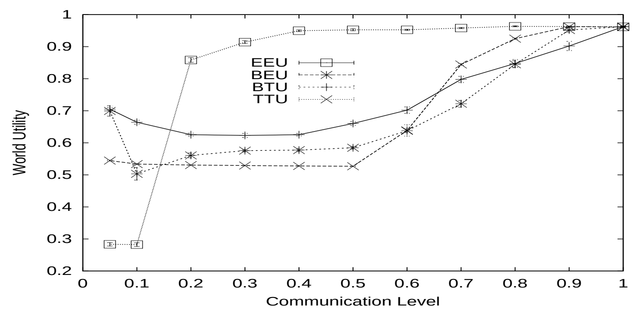

Figure 1: Performance of four utility functions without teams for a range of communication levels. For moderate communication levels EEU performs best. For very low communication BTU performs best since, it uses information from world utility.

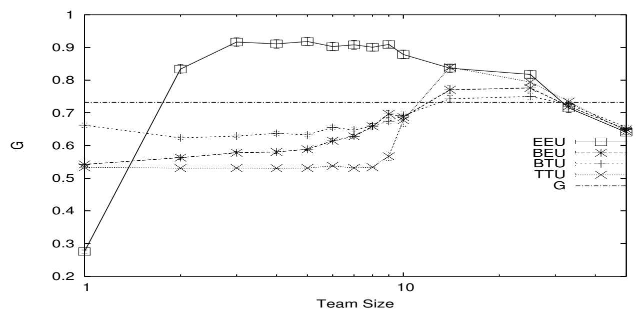

Figure 2: Performance of four utility functions at 10% communication, using teams. EEU performs best for most team sizes under normal learning time.

teams is low. While team information sharing can be seen simply as increasing the communication level, we assume it is added, under the new constrains of team formation, on top of a different communication system with a fixed communication level. Figure 2 shows the tradeoffs between choices of team size at a low level of communication. At most communication levels, there is an optimal team size that lies between the extremes of not having teams (team size = 1), and only having a single team (team size = 100). As the sizes of the teams grow, there is more information sharing, but there is also more noise in each agent's utility, since their utility will be influenced by the actions of more agents. In our problem, the best team size is typically around 5 or 10 agents. This optimum represents to best balance between having small team sizes which produce a more learnable utility and large team sizes which allows for more information sharing.

This work supported by NASA Grant NCC2-5482

3. REFERENCES

[1] J. Fredslund and M. J Mataric. Robots in formation using local information. In *Proc. IAS-7*, March 2002.

[2] D. Pynadath, M. Tambe, N. Chauvat, and L. Cavedon. Toward team-oriented programming. In *Proc. ATAL'99*, pages 77–91, Orlando, Florida, July 1999.

[3] K. Tumer, A. Agogino, and D. Wolpert. Learning sequences of actions in collectives of autonomous agents. In *Proc. AAMAS*, pages 378–385, July 2002.

[4] D. H. Wolpert and K. Tumer. Optimal payoff functions for members of collectives. *Advances in Complex Systems*, 4(2/3):265–279, 2001.

[1] The first two letters of the utility represent how the two terms of the difference utility get their information. "B" stands for "broadcast", "T" stands for "truncated" since the hidden values are just thrown away, and "E" stands for "estimated."