Surrogate Difference Evaluations with Limited Peer to Peer Communications

Main Track Extended Abstract

Steven King
University of Nevada, Reno
Reno, Nevada
kead1@sbcglobal.net

Scott Forer
University of Nevada, Reno
Reno, Nevada
sforer580@gmail.com

Adarsh Kesireddy
University of Nevada, Reno
Reno, Nevada
akesireddy@nevada.unr.edu

Logan Yliniemi
University of Nevada, Reno
Reno, Nevada
logan@unr.edu

ABSTRACT

In this work, we introduce Surrogate Difference Evaluations — an agent-specific credit assignment structure in the vein of Difference Evaluations which only uses local information, and peer-to-peer communicated information. Surrogate Difference Evaluations allow agents to attain 95% performance while operating with 0.04% as much communicable volume.

KEYWORDS

Multiagent Systems; Difference Evaluations; Surrogate Difference Evaluations

ACM Reference Format:
Steven King, Scott Forer, Adarsh Kesireddy, and Logan Yliniemi. 2018. Surrogate Difference Evaluations with Limited Peer to Peer Communications. In *Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), Stockholm, Sweden, July 10–15, 2018,* IFAAMAS, 3 pages.

1 INTRODUCTION

Difference Evaluations are a form of credit assignment used in learning multiagent systems. They have been shown to be effective in coordinating a number of agents toward a shared system-level goal [1]. Difference evaluations have been effectively applied to systems as varied as air traffic control [7], measuring creativity [6], road traffic management [8], and coordinating teams of rovers/agents [5].

However, one common criticism of difference evaluations is that they require global information to calculate. Previous efforts to approximate difference evaluations have relaxed the amount of global information required at a small cost to performance [4], but these methods still require the broadcast of the global evaluation. In the face of limited communications, neither the difference evaluation (which for clarity we term the True Difference Evaluation in this work), nor these approximations are calculable.

Thus, in this work we propose the use of a Surrogate Difference Evaluation, utilizing only local information, and information shared on a peer-to-peer basis. Observations that are known or communicated are used in this calculation, while observations that are neither directly observed nor communicated from direct first- or second-hand knowledge are not. In effect, this is approximating the agent's influence on the portion of the system that it cannot observe or be told about as zero.

In this work, we provide preliminary results showing the efficacy of surrogate difference evaluations in a limited communication environment. We demonstrate that system performance degrades less than 5% in a *limited communication* environment when agents experience a 99.96% reduction in communication volume.

Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), M. Dastani, G. Sukthankar, E. André, S. Koenig (eds.), July 10–15, 2018, Stockholm, Sweden.

2 BACKGROUND

Difference Evaluations are a credit assignment schema for multiagent systems, in which the difference between the performance of the team and performance of the team without that agent's contribution is used as a fitness. G(z) is the Global Evaluation and $G(z_{-i})$ is the Global Evaluation without that agent's contribution. The difference between these two is the Difference Evaluation, which has been denoted D_i in previous works. In this work, to make a strong distinction, we will refer to the **True Difference Evaluation** as TDE_i and our proposed **Surrogate Difference Evaluation** as SDE_i. This True Difference Evaluation is calculated as:

$$TDE_i(z) = G(z) - G(z_{-i}) \tag{1}$$

Approximating Difference Evaluations: Previous efforts have been undertaken to approximate difference evaluations in both discrete domains [4] and continuous domains [2].

In 2015, Colby et al. [4] used the locally available state information and a broadcast of the value of $G(z)$ for each agent to form its own approximation of its difference evaluation. In the discrete domain tested, the agents using a reinforcement learning algorithm were trained on their approximated difference evaluation. These agents converged to similar performance (more slowly) than agents trained using the True Difference Evaluation.

In 2016, this was extended for in continuous state and continuous action domains [2]. In their work, each agent used a neural network to approximate the difference evaluation, and agents used this approximation as fitness in a Cooperative Coevoluation Algorithm (CCEA). The agents in the system only performed 12% worse than the agents in the simulation that used the True Difference Evaluation, while using 90% less state information. These simulations were verified in hardware.

The primary benefit of these approximations is that they did not require each individual agent to have the mathematical form of $G(z)$, but instead relied only on a broadcast of the value of $G(z)$ to form their approximations.

In contrast, the methods proposed in this paper do not even require the broadcast of $G(z)$. We require *no global information*, and instead form our Surrogate Difference Evaluation based on *solely locally available information*, through direct observation or peer-to-peer communication with other nearby agents. Note, however, that SDE depends on the agent's actions (because the actions taken affect the information received) and are thus not a Difference Evaluation by definition.

3 METHOD

Surrogate Global Evaluation

The Surrogate Global Evaluation (*SGE*) is simply the global evaluation ignoring everything that an agent cannot observe. Consider the system vector, z. Of this vector, each agent will only be able to observe a portion of it, and will receive some other information from agents it can communicate with. The union of this information is agent i's known information, which we term $\hat{z}^i$. This will vary from agent to agent. Then, the Surrogate Global Evaluation is defined as $SGE_i = G(\hat{z}^i)$

Surrogate Difference Evaluation

Given the form of a *SGE*, we simply use the form of *TDE* and substitute *SGE* for G:

$$SDE_i(z) = G(\hat{z}^i) - G(\hat{z}^i_{-i}) \tag{2}$$

This Surrogate Difference Evaluation is calculated solely from directly observed or locally peer-to-peer communicated information.

4 DOMAIN & RESULTS

We demonstrate Surrogate Difference Evaluations in an Underwater Multiagent Exploration Domain (UMED). This domain is compelling, because wireless signals are only able to be transmitted through a few meters of water, so global communication is infeasible. The UMED is modeled as a 3-D version of a rover exploration problem [5].

The UMED contains 50 points of interest (POIs). Each POI has a static Cartesian coordinate location (x, y, z) and a value in the range of $\mathbb{R} \in [1, 100]$. Ten agents traverse the 100x100x100 unit domain, following 8 waypoints, observing POIs within their observation radius (10 units), and communicating within their communiciation radius (variable). These ranges are illustrated as circles in **Figure 1**.

We trained a population of 100 policies using a cooperative coevolutionary algorithm (as in [3]) on their individual SDE, though we present results measured on the True Global Evaluation, as this maps to system performance.

Our results (**Figure 2**) show that the communication radius has a very small (less than 5%) impact on performance from full communication down to a communication radius of 5 units, which represents a 99.96% reduction in communicable volume.

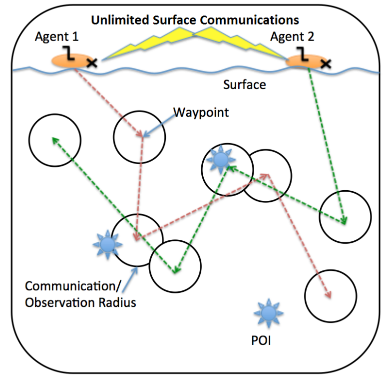

Figure 1: 2-D Cross-Section of UMED Domain Showing Unlimited Communication at the Surface and Limited Communication and Observation Ranges at Waypoints

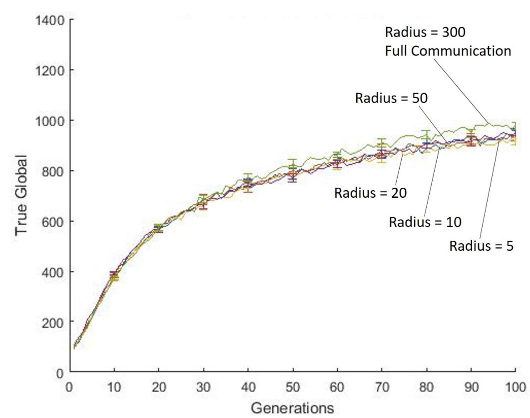

Figure 2: System performance at constant observation and varying communication ranges with agents evolving on Surrogate Difference Evaluations. Reductions in communications radius result in small or statistically insignificant differences; at the minimum communication radius tested (5, vs. 167 required for full communication), system performance degrades less than 5%.

5 CONCLUSION

This work explores the use of Surrogate Difference Evaluations with no global information, instead only using peer-to-peer communication. We show that a Surrogate Difference Evaluation calculated on solely local information could be an effective learning signal: We showed that system performance degrades less than 5% percent while experiencing a 99.96% reduction in communicable volume (full communication vs. communication radius of 5).

Future work on this topic includes testing Surrogate Difference Evaluations in environments where the ability to communicate is not impacted only by proximity, where communication has a cost associated with it, rewarding agents based on their communication behaviors as well as their observation behaviors, and in different types of problem domains.

REFERENCES

[1] A. K. Agogino and K. Tumer. Analyzing and visualizing multiagent rewards in dynamic and stochastic domains. *Autonomous Agents and Multi-Agent Systems*, 17(2):320–338, 2008.

[2] M. Colby, T. Duchow-Pressley, J. J. Chung, and K. Tumer. Local approximation of difference evaluation functions. In *AAMAS 2016*.

[3] M. Colby, L. Yliniemi, and K. Tumer. Autonomous multiagent space exploration with high-level human feedback. *Journal of Aerospace Information Systems*, pages 301–315, 2016.

[4] M. K. Colby, W. Curran, and K. Tumer. Approximating difference evaluations with local information. In *AAMAS*, pages 1659–1660, 2015.

[5] M. Knudson and K. Tumer. Coevolution of heterogeneous multi-robot teams. *Proceedings of the 12th annual conference on Genetic and evolutionary computation*, 2010.

[6] C. Rebhuhn, B. Gilchrist, S. Oman, I. Tumer, R. Stone, and K. Tumer. A multiagent approach to evaluating innovative component selection. In J. S. Gero, editor, *Design, Computing, and Cognition*, 2014.

[7] K. Tumer and A. Agogino. Distributed agent-based air traffic flow management. *AAMAS 2007*.

[8] M. Vasirani and S. Ossowski. A market-inspired approach to reservation-based urban road traffic management. *Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems*, 2009.