An Overview of Coaching with Limitations

Patrick Riley
Computer Science Department
Carnegie Mellon University
5000 Forbes Ave
Pittsburgh, PA 15232
pfr@cs.cmu.edu

Manuela Veloso
Computer Science Department
Carnegie Mellon University
5000 Forbes Ave
Pittsburgh, PA 15232
mmv@cs.cmu.edu

ABSTRACT

As agent relationships become more complex, one challenging relationship that merits study is that of coach or adviser to another agent. Our research on coaching refers to one autonomous agent providing advice to another autonomous agent about how to act. Here we describe the current state of the coach agent framework we are developing. One of our goals is to design a coach agent that can provide advice to many differently built and structured advice taking agents. A coach must then adapt its advice to the capabilities and limitations of the agents it is coaching. This paper explores the effect of advice giving in the presence of limitations.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence; I.2.6 [**Artificial Intelligence**]: Learning

General Terms

Algorithms, Experimentation

Keywords

Coaching, advice

1. INTRODUCTION

In spite of the increasing complexity of the relationships among autonomous agents, one agent coaching or advising another is still somewhat uncommon. We frequently see this relationship among human beings, yet further computational understanding of how this relationship extends to autonomous agents is needed.

As we start to move beyond agent systems of short, limited duration, we believe that advising relationships will become more important. Many current agent systems implicitly allow a complete transfer of knowledge about one agent's behavior structure, both because of the size of the domains and the homogeneity of the agent's representations of the world. In other words, if one agent knows a good way to behave in the world, all agents can easily copy that knowledge. A coach relationship does not assume this is possible. The goal is to obtain knowledge transfer in spite of limitations in communication bandwidth or differences in behavior representation. If the coach wants the other agents to act differently or more optimally, the coach must still consider the best way to *guide* the agents towards this behavior, rather than simply transferring all that knowledge. Additionally, a coach is required to adapt its idea of optimal behavior to the strengths and abilities of the team being coached.

Having a coach agent may make sense in an environment for several reasons. First, an agent may simply have more *a priori* knowledge about the environment, perhaps in the form of a more optimal policy or information about an adversary. A coach agent may also have a better or more complete view of the world, either because of direct sensory availability or because of communication with the other agents. Also, the coach may have more computational resources available to it, allowing it to make better decisions and improve team performance if those decisions can be communicated effectively.

Further, the separation of a coach from an agent team allows a coach to work with different teams of agents. Often, an agent team's performance depends critically on the details of each agent's architecture and behavior decision mechanism. A common coach language should allow a coach to work with many teams. A team is only required to share a common language rather than an entire behavior architecture. For a well structured coach language, support for the language on different behavior architectures should be possible. Any agent team in a sufficiently complex environment will have different strengths and weaknesses and may respond slightly differently to the same advice. A truly effective coach must therefore be able to learn and adapt to the team being coached.

Specifically, our research on coaching refers to one autonomous agent providing advice to another about how to act in the world. Through research on a complex simulated robot soccer domain [2], we have been exploring how one agent can generate advice for teams. With this experience, we have been developing a breakdown of the important components in how a coach agent can effectively advise an agent or agent team.

Figure 1 shows our current framework of the important components of a coach agent. The coach and the coached agents are in an environment (possibly with other agents) where the coach is getting feedback from the agents; our

AAMAS'03, July 14–18, 2003, Melbourne, Australia.

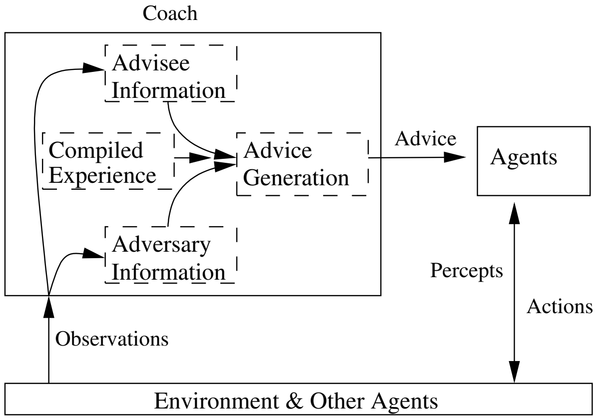

Figure 1: General organization for a coach agent

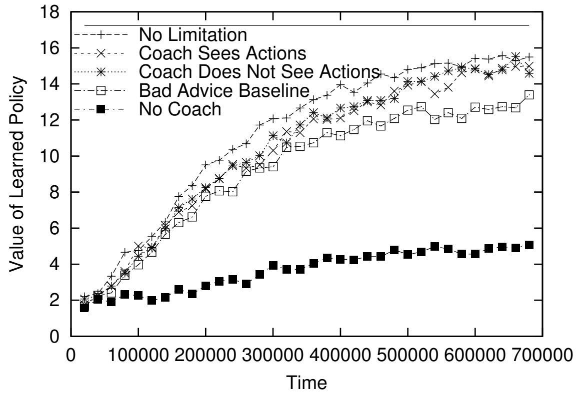

Figure 2: Value of the predators policy for a predator agent learning with a coach advising the agent

goal is not one-shot advice. Note that we are focusing on the coach, or advice-producing agent, not the agents (on the right) which take and use the advice. Our past work deals with generating advice about and in response to adversarial agents in the environment, hence the box for "Adversary Information." While similar coaching techniques may be applicable outside of these adversarial domains, this is one of our primary interests. One important and interesting capability of a coach is the ability to study an environment or recordings of past experiences in an off-line fashion. The "Compiled Experience" box is for summaries of that analysis to provide on-line support.

One hypothesis that underlies our current work is that in order for a coach to be effective coaching multiple different teams of agents, it needs to adapt its advice to peculiarities of the team being coached. The "Advisee Information" box should take observations about the team being coached and use that information to change the advice being sent. This component is the main focus of this paper.

Lastly, the "Advice Generation" box takes input from the others and produces advice in a language appropriately understood by the coached agents.

While our previous work has focused on adapting to an adversary and compiling past experience, the research here deals primarily with adapting advice to the agent being coached. We consider limitations such as limited action abilities of the coached agent, limited communication bandwidth from the coach, limited memory of the coached agent, and whether the coach can see the agent's actions.

2. LEARNING AND LIMITATIONS

We have run a series of experiments in a simulated predator-prey grid world. One predator agent is trying to capture either one of two prey agents that move randomly. Our coach is providing advice to the predator about what actions to take.

One experiment involves the coach providing advice to the predator at every step. The predator is a basic Q-learning agent. However, the agent is limited in the sense that it is unable to perform some of the optimal actions in the environment. The coach is always advising optimal actions.

Figure 2 shows the value of the predator's learned policy over time. The "No Coach" lines shows the agent learning with no coach advice. The "Bad Advice Baseline" shows the coach always advising the action which the agent can not do (which only happens in some of the states). The "No Limitation" line is for the agent running without any actions disabled. The other two lines show how the coach can improve the learning of the agent. This result is similar to, but an extension of previous work by Maclin and Shavlik [1]. They show how the point in learning at which advice is given affects the learning rate. Here, we consider variations on how the advised agent receives advice, and, in forthcoming publications, how limited bandwidth affects the learning.

3. CONCLUSION

Results in our controlled predator-prey show that more advice improves the advisee's performance, even under limitations. This work is part of our ongoing research to understand coach/advice based relationships between automated agents. We plan to further explore the effects of state and action space sizes on the effectiveness of giving advice in controlled environments such as this. Also, while the current advice format has the virtue of simplicity, talking about single states will not scale to larger domains. An advice language must be able to talk about groups of states and actions in order to scale well.

The challenge of coaching is both to learn about an environment and the agents in it and to effectively communicate that information, taking into account the abilities and responses of the receiving agents. This work is one step towards answering that challenge.

4. ACKNOWLEDGMENTS

This research was sponsored in part by the United States Air Force under Grants Nos F30602-00-2-0549 and F30602-98-2-0135 and by an NSF Fellowship. The content of this publication reflects only the position of the authors.

5. REFERENCES

[1] R. Maclin and J. W. Shavlik. Creating advice-taking reinforcement learners. *Machine Learning*, 22:251–282, 1996.

[2] P. Riley and M. Veloso. Planning for distributed execution through use of probabilistic opponent models. In *AIPS-2002*, pages 72–81, 2002.