Multi-Agent Plan Adaptation Using Coordination Patterns in Team Adversarial Games

(Extended Abstract)

Kennard Laviers
U. of Central Florida
klaviers@knights.ucf.edu

ABSTRACT

One issue with learning effective policies in multi-agent adversarial games is that the size of the search space can be prohibitively large when the actions of all the players are considered simultaneously. In most team games, players need to coordinate to accomplish tasks, either in a preplanned or emergent manner. An effective team policy must generate the necessary coordination, yet considering all possibilities for creating coordinating subgroups is computationally infeasible. I propose that reusable coordination patterns can be identified from successful training exemplars and used to guide multi-agent policy search. Experiments are conducted within the Rush 2008 football simulator and show how an analysis of mutual information and workflow can be used to identify subgroups of players that frequently coordinate within a particular formation. Using a K* classifier we devised a system to learns a ranking of the impact of subgroups on offensive performance. Results show how we can use knowledge of the top-ranked subgroup to focus search using two different policy generation methods 1) play adaptation and 2) UCT Monte Carlo (MC) planning. Our method produces superior plans which doubles the offensive team's performance in the Rush 2008 football simulator over prior methods.

Categories and Subject Descriptors

I.2.1 [**Applications and Expert Systems**]: Games

Keywords

opponent modeling, multi-player games, mutual information, play adaptation

1. INTRODUCTION

Effective player coordination has been shown to be an important predictor of team success in adversarial games such as Robocup soccer [9]. Much work has centered on the problem of role allocation, correctly allocating players to roles that are appropriate for their capabilities and smoothly transitioning players between roles [6]. However, in complex team tasks performed by larger teams, roles cannot necessarily be accomplished by a single player. In this case, the team must divide into subgroups to simultaneously handle multiple complex tasks. Unlike agents in coalitions, subgroups within teams share a common global utility metric (team success) but are addressing different tasks or goals. In the worst case, determining which players to group together to accomplish a task requires searching over the partition set of potential team assignments (the Bell number) [8].

My early research introduced a novel method for discovering which agents will make effective subgroups based on an analysis of game data from successful team plays. After extracting the subgroups we employ a supervised learning mechanism to identify the *key group* of players most critical to each play. We evaluated our subgroup extraction method using the Rush 2008 football simulator (RFS). RFS simulates a modified version of American football and was developed from the open source Rush 2005 game.

There are three general types of cues that can be utilized to identify subgroups: spatial, temporal, and coordination dependencies. Our subgroup extraction method uses all of these to build a candidate set of subgroups. By examining mutual information between the offensive player, defensive blocker, and ball location along with the observed ball workflow, we can determine which players frequently coordinate in previously observed plays. Although automatic subgroup identification could be useful for applications such as opponent modeling or game commentary, we demonstrated how extracted subgroups can be used to limit the search space when creating new multi-agent plays using two play generation methods: 1) play adaptation of existing plays and 2) UCT MC planning.

Although there have been other studies examining the problem of recognizing formations in football [2], Robocup [5], and military exercises [7], the problem of identifying subgroups in RFS is novel because these dynamic subgroups are not equivalent to static football formations. In particular, a single formation can lead to a variety of plays, each of which can engender multiple subgroups. We demonstrated that candidate subgroups extracted from observations of successful plays are useful building blocks to generate new plays.

Prior work on finding coordination patterns from an automated analysis of historical play data in team adversarial games falls into four general categories: 1) formation matching based on static spatial patterns, 2) data mining approaches, 3) classification-based play recognizers, and 4) coaching systems. Research on formation matching focuses on matching noisy game data to existing formation templates. Formations are typically known a priori, rather than automatically extracted, and the techniques robustly identify correspondences between the template patterns and the data.

2. RESEARCH AND EXPERIMENTATION

Our approach is to identify subgroups of coordinated players by observing a large number of football plays. In early experiments we demonstrated appropriately changing critical subgroup (e.g., QB, RB, FB) behaviors during an offensive play, in response to a recog-

Cite as: Multi-Agent Plan Adaptation Using Coordination Patterns in Team Adversarial Games (Extended Abstract), Kennard Laviers , *Proc. of 9th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2010)*, van der Hoek, Kaminka, Lespérance, Luck and Sen (eds.), May, 10–14, 2010, Toronto, Canada, pp. XXX-XXX.

nized defensive strategy, significantly improves yardage. We went on to develop a technique to automatically determine critical subgroups of players by an analysis of spatio-temporal logs to find all sub-groups, and supervised learning to determine which groups will garner the best results. Additionally we explored two different techniques for creating new plays: 1) dynamic play adaptation of existing plays, 2) a UCT MC search.

To determine which players should be grouped together dependencies among the eight players for each formation must be understood. All players coordinate to some extent but some players' actions are so tightly coupled that they form a *subgroup* during game play. Changing commands for an athlete in a subgroup without adjusting the others causes the play to lose cohesion. We identify subgroups using a combination of two methods; statistical analysis of player trajectories and workflow.

Mutual information (MI) between two random variables measures their statistical dependence. Inspired by this, our method for identifying subgroups attempts to quantify the degree to which the trajectories of players are coupled, based on a set of observed instances of the given play. However, the naive instantiation of this idea, which simply computes the dependence between player trajectories without considering the game state is doomed to failure. This is because offensive players' motions are dominated by three factors: 1) what the player is supposed to do according to the playbook, 2) position of the ball, and 3) position of the defensive player assigned to block him. So, to calculate the relationships between the offensive players we need to place their trajectories in a context that considers these factors. Rather than computing statistics on raw player trajectory, we derive a feature that includes these factors and compute statistics between the feature vectors.

We compute sets of features from the collection of observed plays for a given pair of offensive players treating observations through time as independent measurements. We model these features as 2D Gaussian distributions with diagonal covariance. We quantify independence between feature distributions using the symmetricized Kullback-Leibler [4] divergence. In addition to finding the MI between player we must also determine relationships based on possession of the football. When the quarterback hands the ball off to the running back or fullback their movements are coordinated for only a brief span of time before the ball is transferred to the next player. Consequently, we also find groups formed by transference of the ball between players and add them to our groups.

Play-adaptation requires identifying the extracted subgroup which will produce the most yardage when changed. To do this we learn a prediction model (K*) for the yardage impact of changing different extracted subgroups.

To adapt the current offensive play we employ a form of dynamic play adaptation. Based on our estimate of the most likely defensive formation sensed early in the play (at t=3), we switch the key subgroup to the play that has the best a priori chance of countering the opponent's strategy. Early experiments demonstrated switching all players after execution has started is *less* effective than adapting the play in a limited manner by only changing the behavior of a key subgroup. To recognize the defensive play in progress we trained a set of support vector machines (SVMs) based on observed player trajectories. We rely upon these to recognize the opponent's strategy at an early stage in the play and we identify the strongest counter based on the yardage history of the offensive playbook against the recognized defense. This can be precomputed, and is therefore an efficient lookup table indexed by the current offensive play and the likely defense.

In addition to play adaptation, we investigated an alternate policy generation method, MC UCT [3], which was successfully demonstrated in real-time strategy games [1]. Using the top-ranked extracted subgroups to focus action investigations yielded significant run-time reduction over a standard Monte Carlo UCT implementation. To search the complete tree without using our subgroup selection method would require an estimated 50 days of processing time as opposed to the 4 days required by our method.

We evaluated this approach on passing plays, which requires tightly coupled coordination between multiple players to succeed. Our version of UCT was seeded with Pro formation variants (4–8). Overall, the UCT and K* plays consistently outperform the baseline Rush system and plays generated using domain knowledge. Furthermore, on average, UCT outperforms the best dynamic play adaptation method by almost 2 yards (30% improvement).

3. CONCLUSION AND FUTURE WORK

A significant challenge in action selection in multi-agent adversarial games is identifying effective player combinations. We demonstrated how to identify and reuse coordination patterns to focus a search over the space of multi-agent policies, without exhaustively searching the set partition of player subgroups. By sorting candidate subgroups using a ranking learned by K*, we reliably identify effective subgroups and improve performance of dynamic play adaptation and the speed of an MC search. In future work, we plan to explore the applicability of these coordination patterns at focusing search in multi-agent reinforcement learning problems.

Acknowledgments

Dr. Gita Sukthankar is the doctoral advisor for this research.

4. REFERENCES

[1] R. Balla and A. Fern. UCT for tactical assault planning in real-time strategy games. In *Proceedings of International Joint Conference on Artficial Intelligence*, 2009.

[2] R. Hess, A. Fern, and E. Mortensen. Mixture-of-parts pictorial structures for objects with variable part sets. In *Proceedings of International Conference on Computer Vision*, 2007.

[3] L. Kocsis and C. Szepesvari. Bandit based monte carlo planning. In *Proceedings of European Conference on Machine Learning*, 2006.

[4] S. Kullback and R. A. Leibler. On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, 1951.

[5] P. Riley, M. Veloso, and G. Kaminka. An empirical study of coaching. In H. Asama, T. Arai, T. Fukuda, and T. Hasegawa, editors, *Distributed Autonomous Robotic Systems 5*. Springer-Verlag, 2002.

[6] P. Stone and M. Veloso. Task decomposition, dynamic role assignment, and low-bandwidth communication for real-time strategic teamwork. *Artificial Intelligence*, 12:pp.241–273, 1999.

[7] G. Sukthankar and K. Sycara. Robust recognition of physical team behaviors using spatio-temporal models. In *Proceedings of International Conference on Autonomous Agents and Multi-agent Systems*, 2006.

[8] G. Sukthankar and K. Sycara. Simultaneous team assignment and behavior recognition from spatio-temporal agent traces. In *Proceedings of National Conference on Artificial Intelligence*, July 2006.

[9] M. Tambe. Towards flexible teamwork. *Journal of Artificial Intelligence Research*, 7:83–124, 1997.