Leveraging Communication for Player Modeling and Cooperative Play

Trevor Sarratt
Advisor: **Arnav Jhala**
{tsarratt, jhala}@soe.ucsc.edu
Computational Cinematics Studio
UC Santa Cruz, Santa Cruz, CA 95064

Abstract

Collaboration between agents and players within games is a ripe area for exploration. As with adversarial AI, collaborative agents are challenged to accurately model players and adapt their behavior accordingly. The task of cooperation, however, allows for communication between teammates that can prove beneficial in coordinating joint actions and plans. Furthermore, we propose extending established multi-agent communication paradigms to include transfer of information pertinent to player models. By querying goal and preference information from a player, an agent can reduce uncertainty in coordination domains, allowing for more effective planning. We discuss the challenges as well as the planned development and evaluation of the system.

Introduction

Virtual agents can take on many roles within games. Adversaries challenge players to learn, adapt, or improve skills. More narrative-oriented agents develop plot through exposition. Likewise, agent teammates provide assistance toward achieving a player's goals. This last type of interaction holds much potential for exploration for game AI researchers.

The task of collaboration in games can be posed in a variety of ways, each with its own relevant work within the AI community. From a purely agent-based perspective, it can be described as a multi-agent system (Niazi and Hussain 2011) with the primary direction of coordinating toward the completion of a goal. Removing any assumptions about a teammate's particular planning approach and belief system sets it within the recent work in ad hoc autonomous agent teams (Stone et al. 2010). The targeting of human-agent teams requires the use of player modeling and, potentially, *theory of mind* concepts regarding what beliefs a player may have about the game state and also regarding his or her companion(s) (Whiten 1991).

Planning-based approaches have found success in recent years in various games. Orkin (2006) demonstrated the capabilities of goal oriented action planning in the game *F.E.A.R.* Hierarchical task networks have similarly been implemented in the game *Killzone 2* (Verweij et al. 2007). Planning serves as an adaptive action control approach in contrast to more static approaches such as behavior trees which are commonly used in games, one such example being *Halo* (Isla 2005).

This paper describes the current direction of work in planning techniques for collaborative agents for application in games. Specifically, we discuss the challenges involved in muti-agent planning under uncertainty of a teammate's goals, values, and planned course of action. Our proposed approach leverages communication to request information from the player in order to more effectively plan for assistive behavior.

It should be noted that this work is still in its preliminary phase, with much of the intended progress to come over the next couple of years. It builds upon a strong, well-established theoretical foundation, however, and we have observed promising results in our explorations into identifying player models and adapting to situations where players switch goals.

Related Work in Team Game AI

Due to the computational contraints within games, cooperative behaviors have often been reduced to simple heuristics or faked altogether. For example, Abraham and McGee (2010) propose a dynamic teammate AI for the game *Capture the Gunner*; however, it depends largely on authored behaviors dependent on where the player is in relation to the teammate and opponent. Repenning (2006) developed a diffusion heuristic for coordinating teams of agents for broad application but relies on tuned factors and authored, static behaviors for coordination. Additionally, some recent work has focused on the extension of common game AI approaches to support player-agent interaction, such as the addition of collaborative actions to behavior queues (Cutumisu and Szafron 2009). A genre-specific categorization of teammate approaches in games can be found in (McGee and Abraham 2010). Furthermore, the authors provide discussion on the relative sparsity of work on the topic, with much left to do in the areas of inference and communication, two key directions for our work.

As McGee and Abraham (2010) note, many AI teammates simply support players by having and acting on the same goal, without perceiving, predicting, or otherwise considering the player's behavior. A functional, if illusory, technique for team AI has been to simply communicate agent actions. Lidén (2003) observes that agent teamates calling out "Taking cover!" while ducking behind a wall or "Flanking" while moving around a corner give the impression of coordinated behaviors, even without any intentional effort to do so. When such teammates are allied with the player, it shifts the responsiblity of planning and executing the joint action, whether it be giving suppressing fire or strafing the opposite side of an enemy, onto the player.

In recent years, a few papers have laid groundwork for truly collaborative agents, complete with goal inference and decision-theoretic action selection. Nguyen et al. (2011) proposed a system of modeling a player's potential goals as states within a Markov decision process (Bellman 1957). This formulation allows for reasoning over possible transitions between goals. As action policies were evaluated by value iteration, the solution technique is rather limited to the size of the state space. Macindoe (2012) implemented a similar system, though adopting a partially observable Markov decision process (POMDP) formulation as well as utilizing a sampling-based solver for improved computation time and scalability.

We leave work in multi-agent systems, communication, modeling and other subjects related to the theoretical foundations of this proposal for the Proposal and Challenges sections of this paper, as needed.

Proposal

We propose a system extending established decision-theoretic communication work within the multi-agent teams domain. A common framework for such problems is the decentralized partially observable Markov decision problem (DEC-POMDP) (Bernstein, Zilberstein, and Immerman 2000). Much of the existing work in DEC-POMDPs restricts communication to observation histories (Roth, Simmons, and Veloso 2006) or sub-goal information (Goldman and Zilberstein 2003). In domains where the agents operate under identical state evaluation and planning, often no further communication is necessary, though suggested actions (Barrett et al. 2014), commitments, and reward/punishment feedback (Goldman and Zilberstein 2003) among teammates are occasionally discussed as additional options.

Specifically, we intend to add functionality for communicating aspects of player models, such as preferences toward goals or types of behavior. The intent of the work is to allow an agent to query information from the player, in a similar manner to (Roth, Simmons, and Veloso 2007), such that an agent teammate can more effectively plan its own policy to support or complement the player's. The broader goal is to explore communication as a strategy for reducing uncertainty in the player's model in addition to inference via observation of behavior.

An initial version would likely restrict the space of goals to simple functions over game state variables, for example minimizing time to completion or maximizing number of coins collected. In addition to inferring the importance of various tasks via observing player actions, an agent could decide to ask the player about his or her feelings toward a state variable, which may be helpful early in a trial when few observations have been made or under the circumstance that the player's behavior has been inconsistent with a potential goal.

Evaluation

Along with implementing a prototype system, our task will be to find a suitable evaluation domain. We will adapt an existing game or design our own. For comprehensive evaluation, the game will need either multiple potential goals or multiple ways to go about a goal. An agent's task, then, will be to infer or query the player for the set of preferences, goals, or other model component and plan accordingly. Initial tests will use simulated humans with a static set of such goals or utility values. Macindoe (2012) evaluated POMCoP in a similar manner, though the game limited the player to one of three choices, which the simulated humans never switched between. We hope to test within a domain that permits the player to have multiple objectives. The space of models comprised of any number of a set of objectives is combinatorial in size, adding to the motivation for communication among teammates.

Once validated with agents, we will begin testing with human players. Many characteristics of play may give insight into the performance and potential deficiencies of the system. Particularly, we are interested in the time to converge to a model, ability to form a coordinated plan, and comparison with human-human teams. As little directly comparable work exists, the system will be compared against itself with communication disabled as an evaluation of the benefit of our communication layer.

Human-human teams may be difficult to analyze as a comparator as a second human player may have his or her own goals and preferences that may be in conflict with his or her teammate. Comparing the performance of a second human against that of an agent whose only goal is to support its teammate may be unfair. We could, of course, give the agent its own utility function for independent goals, but determining the relative importance of the agent's personal goals with respect to aiding the team is still plagued with potential biases. Qualitatively, a questionaire would provide information on how well a player feels his or her teammate, agent or human, inferred his or her goals and adapted its behavior accordingly.

Challenges

The difficulties in building a comprehensive system for collaborating with humans are numerous. Through characterizing these challenges, we intend to illustrate the potential for research in this area. By no means is this list comprehensive, as such a task may entail tackling problems that are rarely or never encountered within games. Instead, we focus primarily on those of direct relevance to our proposed areas of exploration.

Complexity

One of the foremost hurdles for multi-agent team decision problems is computational complexity. Uncertainty on both the player's and the agent's sides with regard to game state and observation history within the game (what has or has not been perceived in the game) in MDP-based formulations of games falls under the category of DEC-POMDPs. Even with a finite horizon assumption for planners, the complexity of finding an optimal joint policy is NEXP-complete (Bernstein, Zilberstein, and Immerman 2000).

Fortunately, it is often possible to reduce the computation required by allowing simplifying assumptions and accepting locally optimal solutions. For example, Nair et al. (Nair et al. 2003) propose fixing teammate policies and searching for locally optimal agent policies until an equilibrium is reached, resulting in a significant reduction in computational time. In the vein of simplifying the problem directly, providing an agent with player action and observation histories, either via the game itself or free communication between agent and player, can allow for scenarios to be posed as single-agent POMDPs (Pynadath and Tambe 2002), which have PSPACE complexity. POMDPs have had considerably more advances than their decentralized counterparts and are frequently solved via dynamic programming (Barto 1998) or sample-based techniques (Silver and Veness 2010).

Recursive Modeling

A large contributor to the complexity of multi-agent planning is recursive modeling, particularly in adversarial scenarios. However, while Tambe (1995) and Gmytrasiewicz and Durfee (1995) note the importance of recursive agent modeling, the assumption of bounded rationality (Simon 1957) can often yield more human-like results (Pynadath and Marsella 2005). At the far end of this extreme, Mundhe and Sen (2000) observed that modeling other agents as being non-recursive and having fixed, probabalistic policies can still lead to convergence to optimal policies in certain domains. We hope to test various restrictions on recursive modeling in order to further reduce the computational requirements of our work.

Learning Player Models

Much work exists on the topic of player modeling. However, many common approaches, such as machine learning, require a bulk of training examples that may require much time to collect in a game. This can be handled by having trained models from a variety of players prior to the instance of play in question. Barrett et al. (2012) further outline a procedure for incorporating current player data into a previously learned model via transfer learning. While this serves as a potential solution to personalized teammate AI, we find undesirable the need for prior data and the complicated nature of blending in new data. For this reason, we prefer to explore communication as a more direct and trustworthy source of player model information.

Moreover, as a point of discussion, while modeling players is a key factor in the success of our approach, the particular models used are less of a focus than the communication aspect of our research. We require only that the models be formed in such a way that specific aspects can be queried, for example by asking questions such as "How do you feel about X?", and that agents should be able to reason over the uncertainty of inclusion or weighting of a component.

Progress

In preparation for the proposed system, our initial work has explored adapting player models to Monte-Carlo tree search, with comparison to value iteration in a partially observable domain (Sarratt, Pynadath, and Jhala In press). Furthermore, various parameters of MCTS were shown to affect belief convergence, which is valuable information for authoring comparable AI systems. Identifying a player's true model from a set of potential models can, however, require many observations of behavior. Planning under the uncertainty of many potential player models is computationally taxing and potentially limiting, given the real-time demands of games. This motivates our communication extensions, as direct discussions of player preferences may quickly reduce the model space, allowing for more effective planning.

Furthermore, we have developed, but not yet submitted for publication, a new approach to adjusting beliefs about player intentions. The new method utilizes the nature of various tasks to enhance belief convergence. Reducing the time for a collaborative agent to identify a change in a player's behavior or intended goal allows for a more responsive change in the agent's policy.

Some of our early exploration in player modeling involved estimating features of play from only very initial data take from a player's behavior. The general idea was to identify similar players using the initial data, then calculate approximations of the missing features using the other players' data until enough data could be collected to create a full model of the player. While the results were promising, the approach required much existing player data and was only useful until information from a variety of actions could be captured. Given that preliminary information in gameplay can be rather noisy, we suspect that communication over preferences will prove to be much clearer oracle of eventual behavior.

Conclusion

In summary, we present a plan of exploration within collaborative game AI. Drawing from many of the established results in related domains, we propose to use communication to further the effectiveness of AI agents in support of players. Agents will complete this specifically by requesting information such that uncertainty in the player's preferences is reduced or eliminated, allowing for more effective planning. With the completion of this project, we anticipate numerous contributions both within the games research community and for the broader AI audience. These include:

- Extension of communication work in multi-agent systems, previously limited to world state information.
- Insight into the utlilty of communication in ad hoc settings.
- Continued progress in collaborative AI for games.
- A novel approach to assistive technologies, with applications in domains such as robotics and human-computer interaction.

References

Abraham, A. T., and McGee, K. 2010. Ai for dynamic team-mate adaptation in games. In *Computational Intelligence and Games (CIG), 2010 IEEE Symposium on*, 419–426. IEEE.

Barrett, S.; Stone, P.; Kraus, S.; and Rosenfeld, A. 2012. Learning teammate models for ad hoc teamwork. In *AAMAS Adaptive Learning Agents (ALA) Workshop*.

Barrett, S.; Agmon, N.; Hazon, N.; Kraus, S.; and Stone, P. 2014. Communicating with unknown teammates. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*, 1433–1434. International Foundation for Autonomous Agents and Multiagent Systems.

Barto, A. G. 1998. *Reinforcement learning: An introduction*. MIT press.

Bellman, R. 1957. A markovian decision process. Technical report, DTIC Document.

Bernstein, D. S.; Zilberstein, S.; and Immerman, N. 2000. The complexity of decentralized control of markov decision processes. In *Proceedings of the Sixteenth conference on Uncertainty in artificial intelligence*, 32–37. Morgan Kaufmann Publishers Inc.

Cutumisu, M., and Szafron, D. 2009. An architecture for game behavior ai: Behavior multi-queues. In *AIIDE*.

Gmytrasiewicz, P. J., and Durfee, E. H. 1995. A rigorous, operational formalization of recursive modeling. In *ICMAS*, 125–132.

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. ACM.

Isla, D. 2005. Handling complexity in the halo 2 ai. In *Game Developers Conference*, volume 12.

Lidén, L. 2003. Artificial stupidity: The art of intentional mistakes. *AI Game Programming Wisdom* 2:41–48.

Macindoe, O.; Kaelbling, L. P.; and Lozano-Pérez, T. 2012. Pomcop: Belief space planning for sidekicks in cooperative games. In *AIIDE*.

McGee, K., and Abraham, A. T. 2010. Real-time team-mate ai in games: A definition, survey, & critique. In *proceedings of the Fifth International Conference on the Foundations of Digital Games*, 124–131. ACM.

Mundhe, M., and Sen, S. 2000. Evaluating concurrent reinforcement learners. In *MultiAgent Systems, 2000. Proceedings. Fourth International Conference on*, 421–422. IEEE.

Nair, R.; Tambe, M.; Yokoo, M.; Pynadath, D.; and Marsella, S. 2003. Taming decentralized pomdps: Towards efficient policy computation for multiagent settings. In *IJCAI*, 705–711.

Nguyen, T.-H. D.; Hsu, D.; Lee, W. S.; Leong, T.-Y.; Kaelbling, L. P.; Lozano-Perez, T.; and Grant, A. H. 2011. Capir: Collaborative action planning with intention recognition. In *AIIDE*.

Niazi, M., and Hussain, A. 2011. Agent-based computing from multi-agent systems to agent-based models: a visual survey. *Scientometrics* 89(2):479–499.

Orkin, J. 2006. Three states and a plan: the ai of fear. In *Game Developers Conference*, volume 2006, 4.

Pynadath, D. V., and Marsella, S. C. 2005. Psychsim: Modeling theory of mind with decision-theoretic agents. In *IJCAI*, volume 5, 1181–1186.

Pynadath, D. V., and Tambe, M. 2002. Multiagent teamwork: Analyzing the optimality and complexity of key theories and models. In *Proceedings of the first international joint conference on Autonomous agents and multiagent systems: part 2*, 873–880. ACM.

Repenning, A. 2006. Excuse me, i need better ai!: employing collaborative diffusion to make game ai child's play. In *Proceedings of the 2006 ACM SIGGRAPH symposium on Videogames*, 169–178. ACM.

Roth, M.; Simmons, R.; and Veloso, M. 2006. What to communicate? execution-time decision in multi-agent pomdps. In *Distributed Autonomous Robotic Systems 7*. Springer. 177–186.

Roth, M.; Simmons, R.; and Veloso, M. 2007. Exploiting factored representations for decentralized execution in multiagent teams. In *Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems*, 72. ACM.

Sarratt, T.; Pynadath, D. V.; and Jhala, A. In press. Converging to a player model in monte-carlo tree search. In *Computational Intelligence and Games, 2014. IEEE Symposium on*.

Silver, D., and Veness, J. 2010. Monte-carlo planning in large pomdps. In *Advances in Neural Information Processing Systems*, 2164–2172.

Simon, H. 1957. A bounded-rationality model of rational choice. *Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting*.

Stone, P.; Kaminka, G. A.; Kraus, S.; Rosenschein, J. S.; et al. 2010. Ad hoc autonomous agent teams: Collaboration without pre-coordination. In *AAAI*.

Tambe, M. 1995. Recursive agent and agent-group tracking in a real-time dynamic environment. In *ICMAS*, 368–375.

Verweij, T.; Schut, M.; Straatman, R.; and Guerrilla, B. 2007. *A hierarchically-layered multiplayer bot system for a first-person shooter*. Ph.D. Dissertation, Masters thesis, Vrije Universiteit of Amsterdam.

Whiten, A. 1991. *Natural theories of mind: Evolution, development and simulation of everyday mindreading*. Basil Blackwell Oxford.