Ad Hoc Coordination in Multiagent Systems with Applications to Human-Machine Interaction

(Doctoral Consortium)

Stefano V. Albrecht
School of Informatics
University of Edinburgh
s.v.albrecht@sms.ed.ac.uk

ABSTRACT

This thesis is concerned with the *ad hoc coordination* problem, in which the goal is to design an autonomous agent which is able to achieve optimal flexibility and efficiency in a multiagent system with no mechanisms for prior behavioural coordination. The thesis is primarily motivated by human-machine interaction problems, which can often be formulated in this setting. This paper gives a brief account of the current state of the thesis and future milestones.

Categories and Subject Descriptors

I.2.11 [**Distributed Artificial Intelligence**]

Keywords

Ad Hoc Coordination; Human-Machine Interaction

1. INTRODUCTION

This thesis is concerned with the *ad hoc coordination* problem. Therein, the goal is to design an autonomous agent, called the *ad hoc agent*, which is able to achieve optimal flexibility and efficiency in a multiagent system that admits no prior coordination between the ad hoc agent and the other agents. *Flexibility* describes the ad hoc agent's ability to solve its task with a variety of other agents in the system. *Efficiency* is the relation between the ad hoc agent's total payoff and time needed to solve the task. *No prior coordination* means that the ad hoc agent does not know ahead of time who the other agents are and how they behave. In particular, there are no prior agreements on information sharing, communication and action protocols, technology standards, etc.

This problem is motivated by the fact that there is a growing number of agents which are employed in an increasing number of areas. Given that a primary goal in agents research is to increase the autonomy and, thus, lifetime of agents, it can be expected that agents based on different technologies may have to interact in nontrivial ways, without knowing a priori who the other agents are. This motivates both the notion of flexibility, since the other agents could be based on any kind of technology, and efficiency, since there may be no time for long learning periods. Human-machine interaction problems can be viewed as an important special case of ad hoc coordination, since humans have extremely variable behaviour (flexibility) and expect agents to be able to interact quickly (efficiency), while there may be no prior description of the human's behaviour (no prior coordination).

Ad hoc coordination is a problem which has been recognised as important only recently, hence prior work is relatively limited, e.g. [6, 8, 11]. However, the assumptions made by the solutions proposed therein imply that they only address certain aspects of the larger problem. For example, in [6, 8] it is assumed that all agents follow complex pre-specified plans which define roles and synchronised action sequences for each role, and in [1, 12, 13] it is assumed that the other agents' behaviours are a priori known and fixed (i.e. they do not learn), and that all agents, including the ad hoc agent, have common payoffs. Furthermore, the problem descriptions in these works are of a procedural nature, associated with the specific tasks considered therein. Thus, there is a need for a formal model of the ad hoc coordination problem.

A related problem is known in game theory as the *incomplete information game*, in which each player has some private information relevant to its decision making. What relates this problem to ad hoc coordination is the fact that no player knows the private information, and hence behaviour, of any other player. Harsanyi [9] introduced *Bayesian games* in which the private information of a player is abstractly represented by its *type*, admitting a solution in the form of the *Bayesian Nash equilibrium*. However, while the notion of private information is useful to describe ad hoc coordination, the learning processes studied therein (e.g. [7, 10]) are not directly applicable, since the focus has traditionally been on equilibrium considerations but not on efficiency.

This thesis is aimed at developing novel models and solutions for the ad hoc coordination problem. The central approach of the thesis is to combine concepts from incomplete information games and optimal control of agent systems. The following sections briefly describe the current state (February 2013) of the thesis and future milestones.

2. CONTRIBUTIONS

Comparative Evaluation of Multiagent Learning Algorithms in Ad Hoc Coordination Problems

As a first step towards understanding the ad hoc coordination problem, I compared five multiagent learning algorithms in a comprehensive set of ad hoc coordination problems, using criteria such as convergence rate, social welfare and fairness,

Appears in: *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2013)*, Ito, Jonker, Gini, and Shehory (eds.), May, 6–10, 2013, Saint Paul, Minnesota, USA.

and equilibrium attainment [3]. The compared algorithms, while representing major approaches in the field (e.g. opponent modelling, policy hill-climbing, and regret-minimisation) were originally developed in a heterogeneous setting, which means that all agents in the system were identical and a priori aware of this fact. The contribution of this work is to show how these algorithms perform in ad hoc coordination problems with possibly heterogeneous settings, which is valuable information for the design of ad hoc agents. The results show that there is, in fact, no clear winner amongst the compared algorithms, each being superior with respect to some criteria but inferior with respect to others.

A Game-theoretic Model and Best-response Learning Method for Ad Hoc Coordination Problems

As was discussed in Section 1, there is currently no formal model and no general solution for the ad hoc coordination problem. I conceptualise this problem formally using a game-theoretic model, called the *stochastic Bayesian game*, in which the behaviour of a player is determined by its private information, or *type* [5]. Based on this model, I derive a solution, called *Harsanyi-Bellman Ad Hoc Coordination* (HBA), which utilises a set of user-defined types in a planning procedure to find optimal actions in the sense of Bayesian Nash equilibrium and Bellman optimal control. I show how HBA can be implemented as a reinforcement learning procedure and evaluate it in a multiagent logistics domain, showing that it is more flexible and efficient than alternative algorithms. I also conducted a human-machine experiment at a public science exhibition in which 427 participants played repeated Prisoner's Dilemma and Rock-Paper-Scissors against HBA and alternative algorithms, showing that HBA achieved equal efficiency and a significantly higher welfare and winning rate.

Exploiting Causality in the Monitoring Task of Partially Observable Markov Decision Processes

A central assumption in my previous two works was that the state of the system is fully observable. However, there are many domains in which the state is not known with certainty due to incomplete and noisy observations. In an early attempt to extend HBA to partially observable domains, I realised that existing methods for monitoring the state of the system were often infeasible in complex real-world domains. Interestingly, these domains often exhibit a causal relation which I refer to as *passivity*. Intuitively, a state variable is passive if it changes its value only if it is the direct target of an action, or if any of the variables that directly affect it change. This insight led to the formulation of a new algorithm, called *Passivity-based Monitoring*, which exploits passivity to accelerate the monitoring task [4]. The idea is to maintain beliefs over individual aspects of the system, and to perform selective updates over these beliefs by exploiting passivity. I show empirically, in both synthetic systems and a real-world system, that this leads to significant performance speed-ups.

3. FUTURE MILESTONES

Hierarchical Specification of Types

The types used in HBA are *complete* in the sense that they specify behaviours for the entire state space of the problem. This is feasible if the problem is small enough so that the behaviours an agent can exhibit are limited. However, complex problems with large state and actions spaces often allow for very complex behaviours (especially if humans are involved), which may make it impracticable to specify complete types. To my knowledge, this has not been directly addressed in game theory or opponent modelling. Therefore, the next milestone of this thesis is a practical solution to this problem. One way to address this might be a hierarchical representation of types, in which behaviour may be specified for regions in the state space, and in which types may be connected in some intelligent way. An interesting question is if such a representation could be learned from data.

Evaluation in Complex Human-Machine System

A long-term goal of this thesis is to study ad hoc coordination problems in complex multiagent systems which involve humans. To this end, I plan to use the internet-based real-time game *Saga Online* [2]. Therein, the players choose to be "Warriors" or "Magicians" and have to defeat each other using a variety of learnable skills. This game has several useful advantages, such as intuitive user interfaces and control mechanisms, and various play modes including a competitive and cooperative mode. Moreover, the game uses complex world dynamics and has extremely large state and action spaces, thus making it a challenging platform for ad hoc coordination. Given the complexity of the game, I expect that complete specifications of types for HBA will be impracticable. Therefore, this game provides a good test platform for the hierarchical type specifications discussed above. Another idea I am pursuing is to use this game to conduct a behavioural *Turing test* in order to evaluate HBA.

4. REFERENCES

[1] N. Agmon and P. Stone. Leading ad hoc agents in joint action settings with multiple teammates. In *11th Autonomous Agents and Multiagent Systems*, 2012.

[2] S. Albrecht. Analysis and correction of the data communication model of a continuous-time distributed application with emphasis on state consistency and fault tolerance (German). B.Sc. Thesis, Technische Universität Darmstadt, 2010.

[3] S. Albrecht and S. Ramamoorthy. Comparative evaluation of MAL algorithms in a diverse set of ad hoc team problems. In *11th Autonomous Agents and Multiagent Systems*, 2012.

[4] S. Albrecht and S. Ramamoorthy. Exploiting causality in the monitoring task of POMDP. Submitted to 29th Conference on Uncertainty in Artificial Intelligence, 2013.

[5] S. Albrecht and S. Ramamoorthy. A game-theoretic model and best-response learning method for ad hoc coordination in multiagent systems (extended abstract). In *12th Autonomous Agents and Multiagent Systems*, 2013.

[6] M. Bowling and P. McCracken. Coordination and adaptation in impromptu teams. In *Proceedings of the National Conference on Artificial Intelligence*, volume 20, page 53, 2005.

[7] E. Dekel, D. Fudenberg, and D. Levine. Learning to play Bayesian games. *Games and Economic Behavior*, 46(2):282–303, 2004.

[8] M. Dias, T. Harris, B. Browning, E. Jones, B. Argall, M. Veloso, A. Stentz, and A. Rudnicky. Dynamically formed human-robot teams performing coordinated tasks. In *AAAI Spring Symposium*, 2006.

[9] J. Harsanyi. Games with incomplete information played by "Bayesian" players. Part I. The basic model. *Management Science*, 14(3):159–182, 1967.

[10] E. Kalai and E. Lehrer. Rational learning leads to Nash equilibrium. *Econometrica*, pages 1019–1045, 1993.

[11] P. Stone, G. Kaminka, S. Kraus, and J. Rosenschein. Ad hoc autonomous agent teams: Collaboration without pre-coordination. In *Conf. on Artificial Intelligence*, 2010.

[12] P. Stone, G. Kaminka, and J. Rosenschein. Leading a best-response teammate in an ad hoc team. In *Agent-Mediated Electronic Commerce: Designing Trading Strategies and Mechanisms for Electronic Markets*, pages 132–146, 2010.

[13] P. Stone and S. Kraus. To teach or not to teach? Decision making under uncertainty in ad hoc teams. In *9th Autonomous Agents and Multiagent Systems*, 2010.