The Formation of Coalitions Among Self-Interested Agents

Steven Ketchpel

Stanford University
Computer Science Department
Stanford CA, 94305
ketchpel@cs.stanford.edu

The Problem

Researchers in the multi-agent systems community of DAI assume that agents will have to interact with others agents that were designed by different designers for different goals. These diverse agents could benefit each other by collaborating, but they will do so only if the resulting deal is beneficial from each agent's point of view. One useful definition of beneficial is that of economic rationality, maximizing the agent's expected payoff in terms of a utility function.

An open problem in this area is to design a protocol that allows a large pool of agents to determine which of the subsets among them can profit by working together. A solution to a coalition problem is a partition of the agents into subsets (**coalitions**), such that each agent in every coalition receives the most utility it can expect.

Objectives

The coalitions that are formed must be **stable** in the sense that none of the agents would leave their current coalitions to form a new one yielding all of the agents in that new coalition a higher utility than they obtain from their previous coalitions. Formally, the agents $a_1,...,a_N$ are divided into a partition P containing coalitions $C_1,...,C_N$ such that every agent is a member of exactly one coalition. The payoff to an agent is a function u(P, a) of both the partition and the agent. For P to be stable, there must not be any other partition P' forming $C'_1,...,C'_N$ such that $\exists C'_i \in P' \; \forall a_j \in C'_i$ with $u(P', a_j) > u(P, a_j)$. If there were such a C'_i, the agents of that coalition would desert their current coalitions and form C'_i.

Since the agents have different abilities, they may be making different contributions to the final outcome. Therefore, splitting the joint reward equally among the included agents might not be equitable. Finally, there are computational considerations. In addition to efficiency, decentralization is desirable, with more robustness in case of node failures and fewer communication bottlenecks.

Methods and Results

Although the idea of coalition formation is relatively new to the field of artificial intelligence, it has been studied by economists working in game theory. Their solutions cannot be directly applied to problems in computer science, however, since different assumptions are made and different phenomena are modeled. One approach to solving the coalition problem is to integrate work from game theory with traditional computational methods. In fact, this approach is being taken by a number of researchers [(Ketchpel 1993), (Shechory & Kraus 1993), (Zlotkin & Rosenschein 1993)]. One greedy algorithm based on the solution to the stable marriage problem (Gusfield & Irving 1989) meets the criteria of being decentralized and efficient, though may yield results which are not stable (Ketchpel 1993). Agents pair off in coalitions that improve their utility the most, then re-enter as a single "agent". The process continues until no new coalitions are formed. A modification to the algorithm (Ketchpel 1994) proposes a "two agent auction" mechanism for cases where the value of collaboration is uncertain. Many questions in this area still need to be addressed. For example, some problems have no stable solution, others have several. Giving the agents more information about each other could introduce strategic behavior that requires more game theoretic analysis.

References

Gusfield, Dan and Irving, Robert W. 1989. *The Stable Marriage Problem: Structure and Algorithms.* Cambridge, MA: MIT Press.

Ketchpel, Steven P. 1993. Coalition Formation Among Autonomous Agents. In Pre-Proceedings of the 5th European Workshop on "Modeling Autonomous Agents in a Multi-Agent World" Supplement. Neuchâtel, Switzerland: Institut d'Informatique et Intelligence artificielle, Université de Neuchâtel. [MAAMAW-93].

Ketchpel, Steven. 1994. Forming Coalitions in the Face of Uncertain Rewards. In Proceedings of the Twelfth National Conference on Artificial Intelligence. Menlo Park, CA: AAAI Press.

Shechory, On and Kraus, Sarit. Coalition Formation Among Autonomous Agents: Strategies and Complexity. In [MAAMAW-93].

Zlotkin, Gilad and Rosenschein, Jeffrey S. 1993a. One, Two, Many: Coalitions in Multi-agent Systems. In [MAAMAW-93].