

Towards Efficient Negotiation Mechanisms for Collaboration*

Timothy Rauenbusch

Division of Engineering and Applied Sciences
Harvard University
Cambridge MA 02138 USA
tim@eecs.harvard.edu

Autonomous agents can achieve more by working together in a team than if each agent acted alone, but disagreements about exactly how to accomplish a group task will arise (e.g., over subtask assignment). The goal of this work is to develop a negotiation mechanism that is suitable for agents involved in a collaboration.

A significant problem in multi-agent systems is that there is no central authority to settle disagreements. Even though a resolution may be globally optimal from the group perspective, an individual agent will not agree to a deal unless that deal is in its own self-interest. For this reason, search techniques that find some globally optimal resolution, such as one that seeks to maximize the sum of the utility of agents in the group (Sandholm 1998) are not directly applicable. For example, even though an outcome that gives \$5 to each robot may maximize a group's gain, a self-interested robot would prefer an outcome of \$6, even if that left another robot with \$1.

Much work in AI on negotiation has adopted methods from the economic field of game theory (Kraus, Wilkenfeld, & Zlotkin 1995; Ephrati & Rosenschein 1991) and has focused on the design of negotiation mechanisms (or protocols) that guarantee *stability*, that is, the incentive for agents to adhere to the strategies prescribed by the mechanism designer.

There is a mismatch between these approaches and the needs of agents negotiating in the context of a collaboration. Game theoretic approaches assume that the evaluation of an agent's preferences is cost-free. Additionally, work based on bargaining assumes common knowledge of agents' preferences. Voting can be applied to more general settings but at the cost of wasting resources (i.e., by an externally imposed tax).

Similarly, familiar auction mechanisms where agents can bid for tasks do not meet the needs of collaborative agents. Arguments for efficiency and stability in auctions assume that there will be a large number of agents bidding on a given item.

We argue that negotiation among collaborators requires a new approach and use the SharedPlans formalism for collaboration (Grosz & Kraus 1999) to identify several important properties of such negotiations which conflict with the assumptions of game-theoretic mechanisms. Specifically, negotiations among collaborators are integrative instead of zero-sum; they are frequent, so resources squandered can quickly accumulate; the evaluation of collaborators' preferences involves costs; and frequently only one agent in a group is capable of performing a given task. Thus, in the setting of collaborative activity, the properties of outcome efficiency, minimization of preference computation, and simplicity are more important than stability.

We describe a new mechanism called Blind Mediation (BM) that can be applied in a collaborative setting. BM implements an anytime search algorithm through the space of possible agreements a group could make. It has the desirable properties of outcome efficiency, minimization of preference computation and simplicity. Simulations have shown that the performance of BM (in terms of long term efficiency and decreased computation cost) is significantly superior to full revelation of preferences and negotiations among people.

References

- Ephrati, E., and Rosenschein, J. S. 1991. The Clarke tax as a consensus mechanism among automated agents. In *Proceedings of the Ninth National Conference on Artificial Intelligence*, 173–178.
- Grosz, B., and Kraus, S. 1999. The evolution of SharedPlans. In Wooldridge, M., and Rao, A., eds., *Foundations and Theories of Rational Agency*, number 14 in Applied Logic Series. The Netherlands: Kluwer Academic Publishers. 227–262.
- Kraus, S.; Wilkenfeld, J.; and Zlotkin, G. 1995. Multiagent negotiation under time constraints. *Artificial Intelligence* 75:297–345.
- Sandholm, T. 1998. Contract types for satisficing task allocation: I theoretical results. In *AAAI 1998 Spring Symposium: Satisficing Models*.

*The research reported in this paper was partially supported by National Science Foundation grants IIS-9978343 and CDA-94-01024, and the DARPA Autonomous Negotiating Teams Program, Contract F30602-99-C-0169.