A Study of Computational and Human Strategies in Revelation Games

(Extended Abstract)

Noam Peled
Gonda Brain Research Center
Bar Ilan University, Israel
+972-3-5317160
noam.peled@live.biu.ac.il

ABSTRACT

This thesis focuses on the design of autonomous agents which can negotiate with people using argumentation strategies. Argumentation is the ability to argue and to persuade another party to accept a desired agreement, to acquire or give information, to coordinate goals and actions and to find and verify evidence [13]. Argumentation is endemic to human interaction. It facilitates knowledge about people's positions, and may improve the final outcome of negotiation [1, 2]. Despite the importance of argumentation within the general framework of negotiations, work on argumentation over the last few years has focused almost exclusively on the context of rational interactions between self-interested, automated agents [6, 7].

Categories and Subject Descriptors

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

General Terms

Experimentation

Keywords

Human-robot/agent interaction, Negotiation

1. INTRODUCTION

Game theory researchers have studied persuasion games since the 1980's [8], but most of the progress has been made in the last few years [3, 5, 9]. In these games, a speaker (e.g., a seller) needs to decide how much information to disclose to the listener (e.g., buyer) in an attempt to encourage the listener to take a specific action (e.g., to buy his goods). Several relevant questions were considered in the context of this limited game. For example, Glazer and Rubinstein [5] studied which rules the listener should use to maximize the likelihood of his accepting the request if, and only if, it is justified, given that the speaker maximizes the probability that his request be accepted. Other researchers tackled the problem of persuasion by studying the use of extensive-form games of perfect information to model argumentation [10, 12]. They used standard backward induction techniques to eliminate dominant strategies and characterized Nash equilibrium strategies for limited cases. Another form of research has applied a mechanism design for abstract argumentation which encourages the agents to reveal their true arguments [4, 11]. To summarize, there are very few previous works on argumentation taking human characterization into account. The theoretical perspective will include a model which will try to predict a human player's strategy. Arguing with people raises challenges for reasons similar to those relating to the development of agents that bargain with people (i.e., people are bounded rational and do not maximize expected utility [1, 2]). We cannot assume that people interacting with an automated agent will follow a predefined algorithm for producing argumentation, use equilibrium strategies or even that they will follow a predefined protocol for the argumentation. To the best of our knowledge, there are no systems that can argue with people or provide argumentation when negotiating with or facilitating negotiation between people.

Cite as: A Study of Computational and Human Strategies in Revelation Games (Extended Abstract), Peled Noam, *Proc. of 10th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2011)*, Tumer, Yolum, Sonenberg and Stone (eds.), May, 2–6, 2011, Taipei, Taiwan, pp. 1363-1364.

2. EQUILIBRIUM STRATEGIES

In the first section of the thesis we tackled the following challenges: first, to determine how well computer agents negotiate with people in revelation games where agents use equilibrium strategies that entail deciding whether or not to reveal private information; second, to understand how people relate to agents in such games. We used two types of revelation games that varied the dependencies between players. Each game included a revelation choice followed by two rounds of negotiation. We compared people's performance when playing these games with other people to that of computer agents playing against people. The computer agents used one of two types of possible equilibrium strategies. One type did not reveal its preferences at all during any point the negotiation, while the other type revealed its true preferences at the onset of the negotiation process. Both equilibrium types made competitive, more selfish offers in the first negotiation round and more generous offers in the last round. Depending on their strategy, some agents asked for more resources than they needed if their preferences weren't known. The results of our experiments show that (1) an agent's performance depended on whether they were the last party to make a proposal, but did not depend on whether or not they decided to reveal their true preferences. For people, this trend was reversed. In particular, preference revelation increased the likelihood of agreement for people, but not for agents. (2) Agents performed as well as people when they were the last party to make a proposal, but overall, they were significantly outperformed by people. We conjectured this was because people were reluctant to accept the competitive offers made by agents in the last round. These results thus indicate that preference revelation has a significant positive effect on people's performance but this benefit does not carry over to equilibrium-playing agents when they make strategic-type offers. These results provide insight into people's strategies in revelation games that will facilitate future agent-design in these settings.

3. DECISION THEORY

In the second section, we built a new agent-design that uses a decision-theory approach to negotiating proficiently with people in revelation games. The agent explicitly reasons about the social factors that affect people's decisions whether to reveal private information, as well as the effects of people's revelation decisions regarding their negotiation behavior. It combines a prediction model of people's behavior in the game with a decision-theory approach for making optimal decisions. The parameters of this model were estimated from data about human play. The agent was evaluated playing against both new people and an agent using equilibrium strategies in a revelation game that varied the dependency relationships between players. The results showed that the agent was able to outperform human players as well as the equilibrium agent. It learned to make offers that were significantly more beneficial to people than the offers made by other people while not compromising its own benefit, and was able to reach agreement significantly more often than did people as well as the equilibrium agent. In particular, it was able to exploit people's tendency to agree to offers that are beneficial to the agent if people revealed information at the onset of the negotiation. The contributions of our work are fourfold. First, it formally presents revelation games as a new type of interaction which supports the controlled revelation of private information. Second, it presents a model of human behavior that explicitly reasons about the social factors that affect people's negotiation behavior, as well as the effects of players' revelation decisions on people's negotiation behavior. Third, it incorporates this model into a decision-making paradigm for an agent that uses the model to make optimal decisions in revelation games. Lastly, it provides an empirical analysis of this agent, showing that the agent is able to outperform people and more likely to reach an agreement than people.

4. FUTURE WORK

For future work we have several directions: (a) First, we intend to build an agent who plays revelation games, including more complex argumentation domains. One possibility is a domain where the players are not exposed to each other's resources, and in each negotiation phase they can reveal a subset of their resources. (b) We want to investigate cooperative game theory concepts in the domain of revelation games. According to our intuition, agents playing according to these concepts can play much better against people than against equilibrium agents, mainly because their strategy will be more similar to a person's strategy while playing these games. (c) We want to expand our decision-theory model to be able to grasp the diversity of peoples' social preferences and find distinctive clusters in these preferences. (d) We want to let the agent be exposed to the human player's brain activity while they are playing revelation games, and to use feature-detection algorithms in order to build a prediction model for human strategy based on their brain activity. In this way an agent can learn from past games and can adapt to its opponent while playing.

5. REFERENCES

[1] O. Andersson, M. M. Galizzi, T. Hoppe, S. Kranz, K. der Wiel, and E. Wengstr\öm. Persuasion in experimental ultimatum games. *Economics Letters*, 108(1):16–18, 2010.

[2] T. Ellingsen and M. Johannesson. Anticipated verbal feedback induces altruistic behavior. *Evolution and Human Behavior*, 2007.

[3] F. Forges and F. Koessler. Long persuasion games. *Journal of Economic Theory*, 143(1):1–35, 2008.

[4] J. Glazer and A. Rubinstein. Debates and decisions: On a rationale of argumentation rules. *Games and Economic Behavior*, 36(2):158–173, 2001.

[5] J. Glazer and A. Rubinstein. A study in the pragmatics of persuasion: a game theoretical approach. *Theoretical Economics*, 1(4):395–410, 2006.

[6] N. R. Jennings, P. Faratin, A. R. Lomuscio, S. Parsons, M. J. Wooldridge, and C. Sierra. Automated negotiation: prospects, methods and challenges. *Group Decision and Negotiation*, 10(2):199–215, 2001.

[7] S. Kraus, K. Sycara, and A. Evenchik. Reaching agreements through argumentation: a logical model and implementation* 1. *Artificial Intelligence*, 104(1-2):1–69, 1998.

[8] P. Milgrom and J. Roberts. Relying on the information of interested parties. *The RAND Journal of Economics*, pages 18–32, 1986.

[9] E. T. Pillai. *Two-Sided uncertainty in persuasion games*. PhD thesis, University of Minnesota, May 2009.

[10] A. D. Procaccia and J. S. Rosenschein. Extensive-form argumentation games. In *Proceedings of the Third European Workshop on Multi-Agent Systems (EUMAS-05), Brussels, Belgium*, pages 312–322, 2005.

[11] I. Rahwan and K. Larson. Argumentation and game theory. *Argumentation in Artificial Intelligence*, pages 321–339, 2009.

[12] R. Riveret, H. Prakken, A. Rotolo, and G. Sartor. Heuristics in argumentation: a game-theoretical investigation. In *Proceeding of the 2008 conference on Computational Models of Argument: Proceedings of COMMA 2008*, pages 324–335, 2008.

[13] S. E. Toulmin. *The uses of argument*. Cambridge Univ Pr, 2003.