Understanding How People Design Trading Agents over Time*

(Short Paper)

Efrat Manisterski, Raz Lin and Sarit Kraus†
Computer Science Department
Bar-Ilan University
Ramat-Gan, Israel
{maniste,linraz,sarit}@cs.biu.ac.il

ABSTRACT

As computerized agents are becoming more and more common, e-commerce becomes a major candidate for incorporation of automated agents. Thus, it is vital to understand how people design agents for online markets and how their design changes over time. This, in turn, will enable better design of agents for these environments. We focus on the design of trading agents for bilateral negotiations with unenforceable agreements. In order to simulate this environment we conducted an experiment with human subjects who were asked to design agents for a resource allocation game. The subjects' agents participated in several tournaments against each other and were given the opportunity to improve their agents based on their performance in previous tournaments. Our results show that, indeed, most subjects modified their agents' strategic behavior with the prospect of improving the performance of their agents, yet their average score significantly decreased throughout the tournaments and became closer to the equilibrium agents' score. In particular, the subjects modified their agents to break more agreements throughout the tournaments. In addition, the subjects increased their means of protection against deceiving agents.

Categories and Subject Descriptors

I.2.1 [**Artificial Intelligence**]: Applications and Expert Systems—*Games*

General Terms

Experimentation

Keywords

trading agents, negotiation, equilibrium

1. INTRODUCTION

With the growth and accessibility of the internet and the web, e-commerce has become widely used [5]. As trading in many markets (e.g. NASDAQ) has now become fully electronic, designing automated trading agents has received growing attention. Instead of personally trading online, the trader can now use a computerized agent which trades on his behalf and makes the decisions for him. For example, a person that wants to purchase an item can be represented by a software agent. This agent negotiates for the item on behalf of the person and can eventually purchase the item for him when all conditions are met [9]. Thus, there is a growing importance of understanding how people develop agents for these environments. Indeed, examining the ways in which people design their agents has been established as a key goal in several AI studies [3, 12, 13].

In this paper we present a novel research, which further enhances the understanding of strategic behavior of agents designed by humans. More specifically, we investigate the change in the agents design over time. We focus on the design of trading agents for bilateral negotiations with unenforceable agreements (e.g., as in eBay). We term these agents Peer Designed Agents (PDAs). In this context, it is important to understand how people modify the strategy behavior of their agents, based on performance in the past, as automated agents are, in nature, used in recurring events and transactions. Thus, our work provides a significant contribution in this respect.

In order to simulate a common real-life bilateral negotiation environment we used the domain of a resource allocation game. In the resource allocation game, two sides are given an initial set of resources and a goal. Both sides negotiate the exchange of resources with each other in order for each to achieve its goal. The player's negotiation policy is always accompanied by the decision whether to send the resources which he has agreed upon. If the number of interactions between the players is finite, the equilibrium strategy requires that no exchanges are made. The strategy space in the resource allocation game is richer and larger than most models previously studied in economic research [1, 2, 4, 7, 8]. As a matter of fact, this game supplies a general negotiation platform which is more similar to real-life negotiations than typical economic games.

We ran experiments in which two groups of graduate students were assigned to design agents that play the resource allocation game. Each student was responsible for designing his own agent. Each group received a different version of the resource allocation game that differs in the dependencies between players (the resources the player needs and its dependency on the other player to supply those resources). Each group was involved in several tournaments where each subject's agent played against all other agents in its group, including itself. After each tournament, subjects were permitted to improve their agents based on the feedback they received about their agent's performance in previous tournaments.

*This work is supported in part by NSF under grant #IS0705587 and ISF under grant #1357/07

†Sarit Kraus is also affiliated with the University of Maryland Institute of Advanced Computer Studies.

Cite as: Understanding How People Design Trading Agents over Time (Short Paper), Efrat Manisterski, Raz Lin and Sarit Kraus, *Proc. of 7th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2008)*, Padgham, Parkes, Müller and Parsons (eds.), May, 12-16., 2008, Estoril,Portugal,pp. 1593-1596.

Our results show that the subjects fundamentally modified their agents' strategic behavior throughout the tournaments. Specifically, the agents' average score significantly decreased throughout the tournaments, while in the last tournament the agents' average score was closer to the equilibrium agent's score than in the first tournament. However, even in the last tournament most subjects did not develop agents that adhere to the equilibrium strategy (that is, not sending any resources). The decrease in the agents' average scores can be explained by analyzing their strategies. More agents were less cooperative and their reliability level decreased (i.e., agents did not live up to their promises) throughout the tournaments.

This paper contributes to research on agents' designs in several ways. First, it provides insight into the considerations people take into account when designing agents. Second, the results of these experiments provide information about the changes in agents' behavior throughout the tournaments. These findings can help better understand the agents' behavior over time. These results suggest that the number of tournaments has substantial influence on the market and on agent design.

The rest of the paper is organized as follows. We begin by reviewing related work in Section 2. Then we continue with a description of our trading framework in Section 3. The experiment design is provided in Section 4 and the results in Section 5. Finally, we conclude and discuss future work in Section 6.

2. RELATED WORK

In this paper we explore the strategic behavior of Peer Designed Agents. A very known competition which has shed light on the automated trading agents' strategies is the Trading Agent Competition (TAC) [6, 12, 13]. In this competition entrants develop travel agents that have to arrange itineraries for a group of clients who want to fly from one city to another within a certain time period. There are several fundamental differences between the TAC competition and our work. First, our trading domain is different from the TAC as we focus on bilateral negotiations with unenforceable agreements while TAC focuses on agents' bidding strategies for complementary and substitutable goods. Second, the TAC is a competition in which the agent's target is to win (i.e. to attain the highest number of points/money) while in our work, as in many real life domains, the goal is to accumulate as many points as possible, regardless of the other agents' performance.

Some research in economics has explored the strategic behavior when people are able to revise their strategies based on their performance in the past. Most of them used a strategy method[1] [11]. The strategy method was first proposed by Selten [10]. Similar to developing agents, using the strategy method requires subjects to specify their choices for all information sets of the game and not only the ones that occur during the course of a play of a game. The strategy method is usually used to elicit subjects' strategies in games that are relatively simple (e.g. public goods, Prisoner's Dilemma, ultimatum game and generation game) in contrast to our game which has a richer strategy space and is much more complicated.

3. THE TRADING FRAMEWORK

In order to simulate a general trading framework we have designed the resource allocation game, used as a test-bed to represent various situations in real economic markets.

[1]A strategy method is a known economic experimental methodology which requires people to elicit their actions in every decision node in the game.

We begin by describing the environment description and then we continue on to analyze the different strategies in our framework.

3.1 Environment Description

Each player $i \in \{1,2\}$, is allocated an initial pool of resources R_i^{init}, which are attributed to several types. The goal of the game is to possess a specified set of resources G_i, which includes a certain quantity (zero or more) of each resource type. There are enough resources for both players to satisfy their goals, i.e. $G_1 \cup G_2 \subseteq R_1^{init} \cup R_2^{init}$. However, some of the resources needed by one player may be in the possession of the other. The negotiation protocol consists of a finite number of rounds n. In each round a different agent proposes an offer, while the other agent can respond to it. Each round $0 \leq l \leq n$ is comprised of two phases: a negotiation phase and an exchange phase. In the negotiation phase one of the players makes an offer to exchange resources, $O^l = (Ogive_i^l, Ogive_j^l)$, in which the proposer (player i) promises to send $Ogive_i^l$ resources to the receiver (player j) and in return requests that player j will send $Ogive_j^l$ back to him. Player j should inform the player i whether he accepts or rejects the offer. Afterwards, there is an exchange phase, in which the two players, players i and j, send a set of resources S_i^l and S_j^l, respectively, to the other player. Since agreements are not enforced, each player can break agreements, thus S_i^l and S_j^l can differ from $Ogive_i^l$ and $Ogive_j^l$, respectively. The exchange is executed simultaneously, so the players cannot know in advance whether the opponent will keep his promise. The performance of each player is determined by his score at the end of the game. The score of a player takes into account both the number of resources the player possesses, as well as whether or not he has reached the goal. For each resource the player possesses at the end of the game, the player will receive a score of $Score_{Res}$. In addition, if the player holds his whole target set, he will receive an additional score of $Score_{goal}$. Formally

$$score_i = \begin{cases} Score_{goal} + |R_i^{end}| Score_{Res} & G_i \subseteq R_i^{end} \\ |R_i^{end}| Score_{Res} & \text{otherwise} \end{cases}$$

In our experiment we used $n = 10$, $Score_{goal} = 200$, $Score_{Res} = 10$, $|G_i| = 3$ and $|R_{init}^i| = 8$. Thus, obtaining the target set becomes the most important component of the scoring function.

To generalize and strengthen our results, we used two distinct configurations: the one independent player configuration (OIP) and the asymmetric depth configuration (AD). In the OIP configuration, while one player is *independent* and initially obtains all resources needed for him to reach the goal, the other player lacks two specific resources, and he is *dependent* on the other player's resources. Both players also have some extra available resources, which they do not need to attain their respective goals, and can be used for negotiation purposes. This configuration enables an examination of situations in which the equilibrium strategy is not Pareto-optimal, and cooperation between the players yields better results for both parties: the dependent player can obtain the resources he needs to complete the target set, and the independent player can increase the number of resources he possesses. This setting enables the examination of situations, where one side might gain substantially more from the transaction than the other side. These situations are very common in real life. Consider, for instance, a researcher who crucially needs to buy some books for an important research. While the seller of the books will gain some money from the transaction, the researcher, might gain much more from this deal.

In the AD configuration each of the players needs at least one resource from the other player in order to complete its target set. However, the needs of these players are asymmetric. The first player, the *2-resources player*, needs two resources from the second player, while the second player, the *1-resource player*, needs only one resource from the other player in order to complete its target set. Again cooperation between the players yields better results for both parties: both players can obtain the resources they need in order to complete the target set. This setting enables the examination of a very common situation, where each of the traders significantly gains from the exchange of resources, yet one side is more dependent than the other. Consider, for instance, two researchers who each possess the books that the other needs for his/her research, where each of them needs a different amount of books from the other. In this case both can gain from the exchange.

In the following section we describe the experiment's design and methodology.

4. EXPERIMENT DESIGN

The experiments involved 32 different agents developed by 32 computer science graduate students. The students were divided into two groups of size 15 and 17. All subjects were instructed to design an agent that plays the resource allocation game, which was explained in class. The experiment was identical for the two groups, except that each group had received a different configuration, as described above. Subjects were explained that their agent would play in a tournament against all the other subjects' agents including their own agent in both roles. We kept the identity of all agents antonymous so that agents would not be able to treat any other agent based on its history.

A skeleton of the agent and a support environment were provided to the students. In addition to programming the agents, the subjects were instructed to submit documentation explaining their strategy.

After a given tournament each of the subjects had received feedback about his agent's performance during the tournament. More specifically, each of the subjects received a log file for each game that its agent played, which included the course of the game (that is, the offers that were proposed, the responses to these offers, and the resources that each of the players sent in each round of the game). Then, subjects had the opportunity to revise their agents and resubmit them for the next tournament. We repeated this process three times, so each of the agents participated in four tournaments.

The subjects were motivated to perform well by receiving a course grade. We emphasized that the grade is based only upon the subject's own score, and not upon other subjects' scores. This is similar to real negotiation environments, where traders gain money only according to their own agents' performance. To motivate subjects to make efforts also in the later tournaments, we explained to them that the grade was calculated according to the score obtained by their agents accumulated over all tournaments, whereby later tournaments received much larger weight. Moreover, in order to ensure that in each tournament each of the subjects will do their best and will not count on future tournaments to improve their agents, we did not reveal to them in advance whether there would be additional tournaments or not. In this manner the subjects tended to believe that each tournament was their last one.

5. EXPERIMENTAL RESULTS

The agents participated in four tournaments. When analyzing the performance of the agents we used two benchmark scores. The first benchmark is the score obtained by equilibrium agents. As these agents do not send any resources, their final scores are equal to their initial scores. More specifically, given the OIP configuration, the equilibrium agent's score when it plays the dependent role, denoted by sc_{dep}, equals 80 (8 resources multiplied only by $Score_{Res}$, as it does not complete its target set). The equilibrium agent's scores when it plays both roles in the AD configuration, denoted sc_{2-res} and sc_{1-res} also equal 80, based on the same considerations. On the other hand, the equilibrium agent's score when it plays the independent role, denoted sc_{ind}, equals 280 (it receives $Score_{Goal}$ as it completes its target set, with an additional score of the resources it has, that is 8 resources multiplied by $Score_{Res}$).

The second benchmark is supplied by agents that achieve a Pareto-optimal solution, which is mutually beneficial. In other words, both agents' score at the end of the game is higher than the score they start with at the beginning of the game. Moreover, no agent can increase its score at the end of the game without decreasing the other agent's score. More specifically, in the OIP configuration the dependent agent will complete its target set at the end of the game and the independent player will increase the number of the resources it obtains (it will possess more than 8 resources at the end of the game). This means that the independent player's score will be greater than 280 (as it will possess more than 8 resources at the end of the game) while the dependent score will be greater than 230 (as it will possess at least the three resources that comprise its target set at the end the game). As both players at the end of the game complete their target set and together obtain 16 resources (the only change during the game is the distribution of the resources between the players) the total score of the agents playing both roles at the end of the game equals 560 thus $sc_{dep} = 560 - sc_{ind}$. This implies that the average score, denoted $sc_{OIP-avg}$, that an agent obtains when it plays both roles equals 280. Similar to the OIP configuration, in the AD configuration both players possess at least the three resources that they need to complete their goal which implies that both players' score is greater than 230. As the two agents complete their target set at the end of the game it follows that $sc_{2-res} + sc_{1-res} = 560$ and the average score, denoted sc_{AD-avg} that the agent obtains when it plays both roles equals 280. However, in contrast to the OIP configuration, any distribution of the resources between the agents in which each of the agents possesses its target set at the end of the game is possible.

First we will examine the change in the agents' average score for all the tournaments. When comparing the scores to the benchmarks we observed that the PDA's obtained, on average, higher scores than those of the equilibrium agents. On the other hand, this score is lower than that of the Pareto-optimal agents. Moreover, the PDA's average score decreases with the tournaments and in the last tournament the agent's average score is closer to that of the equilibrium agents' score than in the first tournament. This decrease can be explained by the percentage of agents that complete the target set.

When we analyzed the percentage of agents (not including agents playing the independent role) that completed their target set we saw that that in all considered roles the percentage of agents that complete their target set monotonously decreases over the tournaments. We have observed a significant decrease in the percentage of the dependent agents that completed their target set from the first (55%) to the second tournament (30%) (χ^2, $p < 0.01$). In addition, for all the considered agents, a significant decrease in the percentage of agents that complete their target set is observed from the first tournament to the third tournament (χ^2, $p < 0.05$). This decrease is more pronounced when comparing the first tournament to the fourth tournament (χ^2, $p < 0.01$).

Two possible reasons can explain the decrease in the scores and the percentage of agents reaching the goal. The first is the fact that the number of agreements that were not fulfilled or were partially fulfilled increased. This means that less resources were sent, including resources the agents needed in order to complete their target set. The second explanation is that fewer agreements were reached due to the hardening of the negotiation policy. For example, consider two agents that fulfill their agreements when playing the resource allocation game in the OIP configuration. The first agent, playing the independent role, requires at least three resources from the other agent for each of the resources the other agent needs to complete its target set. However, the other agent playing the dependent role agrees to give at most two resources in return for each resource it needs. Thus the dependent player will never complete its target set as no agreement will be reached.

However, after reviewing the PDAs code and analyzing their strategies, we can deduce that the first explanation is more likely. Not only did most of the subjects not harden their negotiation policy, some of them even softened their negotiation policy. Still, agreements were not fulfilled (or were only partially fulfilled) throughout the tournaments. This can be explained by several reasons. The first can be attributed to the agents' reliability level which decreased throughout the tournaments. Thus, the agent broke more agreements regardless of the other agent's reliability (i.e., the agents did not fulfill their agreements even if the other agents sent all the resources that were agreed upon). The second reason is that agents tended to break more agreements in order to protect themselves against deceivers (i.e., the agents broke their agreements as a result of the other agents' behavior). This preventive behavior is based on deducing the nature of the opponent from previous rounds.

6. CONCLUSIONS AND FUTURE WORK

In this paper we explored the modifications of the strategic behavior of Peer Designed Agents. We focused on bilateral negotiations with unenforceable agreements over time. In order to simulate such environments we used a resource allocation game. This game has a richer strategy space than most of the standard economic games [1, 2, 4, 7, 8]. We tested two different configurations of this game reflecting real life and common situations, which differ from each other by the number of resources each player needs and its dependency on the other player.

Even though the basic configurations are different, the results of our experiment show that agentsŠ strategic behavior was fundamentally modified throughout tournaments. Our experiments showed that both the average score and the percentage of agents that completed their target set decreased significantly throughout the tournaments. This is despite the fact that most of the subjects improved their agents as compared to their previous versions. In both groups the agents become less reliable throughout the tournaments and less likely to keep their agreements regardless of the other agent's behavior.

Moreover subjects learned from previous tournaments and increased their protection against low reliability agents. This behavior can also explain the decrease in the agents' average score since low reliability agents complete their target sets in the last tournament less than in the first tournament due to an increase in the agents' protection level.

Our findings play an important role in understanding dynamic markets in which traders are able to modify their trading agents. Moreover these results have great implications on the agent's design. Thus we recommend that agent designers take the tournament number into account when designing agents for this type of market.

In future work we will investigate how people develop agents in more complex games as well as the behaviors that emerge in online games.

7. REFERENCES

[1] J. Andreoni. Why free-ride? strategies and learning in public goods experiments. *Journal of Public Economics*, 37:291–304, 1988.

[2] C. F. Camerer and E. Fehr. *Foundations of Human Sociality*, chapter Measuring Social Norms and Preferences using Experimental Games: A Guide for Social Scientists. Oxford University Press, 2004.

[3] B. Grosz, S. Kraus, S. Talman, B. Stossel, and M. Havlin. The influence of social dependencies on decision-making: Initial investigations with a new game. In *AAMAS04*, pages 782–789, 2004.

[4] W. Guth and E. Damme. Information, strategic behavior and fairness in ultimatum bargaining: An experimental study. *Journal of Mathematical Psychology*, 42(2):227–247, 1998.

[5] Y. Ishihara, R. Huang, and J. Ma. A real trading model based price negotiation agents. In *AINA*, pages 597–604, 2006.

[6] P. L. Lanzi and A. Strada. Analysis of the trading agent competition 2001. *SIGecom*, 3(2):1–8, 2002.

[7] T. Offerman, J. Potters, and H. A. A. Verbon. Cooperation in an overlapping generations experiment. *Games and Economic Behavior*, 36(2):264–275, 2001.

[8] C. O'Riordan. Iterated prisoner's dilemma: A review. Technical report, Department of Information Technology, NUI, Galway, 2001.

[9] B. Schnizler, S. Luckner, and C. Weinhardt. Automated trading across e-market boundaries. In *Group Decision and Negotiation*, pages 199–201, June 2006.

[10] R. Selten. Die strategiemethode zur erforschung des eingeschränkt rationalen verhaltens im rahmen eines oligopolexperiments. *Beiträge zur experimentellen Wirtschaftsforschung*, 1:136–168, 1967.

[11] R. Selten, M. Mitzkewitz, and G. R. Uhlich. Duopoly strategies programmed by experienced players. *Econometrica*, 65:517–555, 1997.

[12] P. Stone and A. Greenwald. The first international trading agent competition: autonomous bidding agents. *Electronic Commerce Research*, 5(2):65–229, 2005.

[13] M. P. Wellman, A. Greenwald, P. Stone, and P. R. Wurman. The 2001 trading agent competition. *Electronic Markets*, 13(1):4–12, 2003.