The Development of the Strategic Behavior of Peer Designed Agents*

Efrat Manistersky[1], Raz Lin[1] and Sarit Kraus[1,2]

[1] Department of Computer Science
Bar-Ilan University
Ramat-Gan, Israel 52900
{linraz,sarit}@cs.biu.ac.il
[2] Institute for Advanced Computer Studies
University of Maryland
College Park, MD 20742 USA

Abstract. As computerized agents have become more and more common, e-commerce has become 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 a 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. However, 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, while also increasing their means of protection against deceiving agents.

1 Introduction

With the growth and accessibility of the internet and the web, e-commerce has become widely used [13]. As trading in many markets (e.g. NASDAQ) has now become fully electronic [8], designing automated trading agents has received growing attention. Instead of personally trading online, the trader can now use a computerized agent which trades on her behalf and makes the decisions for her. For example, a person who desires 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 her when all conditions are met [21]. Thus, there is growing importance in 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 [10, 25, 26].

Apparently, we could have applied game theory considerations to predict the strategies of these kinds of agents. However, extensive research conducted in behavioral economics and psychology has shown that people play differently from the game theoretic predictions [9, 20]. Beyond the rational utility maximization, humans are also motivated by social factors such as equality, reciprocity, and fairness [6]. Preliminary research on how computer agents are designed by humans has shown that people design agents differently from the way they play. Moreover, they do not use game theory or equilibrium considerations in their design [10].

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, by nature, used in recurring events and transactions. Thus, in this respect our work provides a significant contribution.

In order to simulate a common real-life bilateral negotiation environment we used the domain of a resource allocation game [17]. 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 her goal. The player's negotiation policy is always accompanied by the decision whether to send the resources which she 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, 5, 11, 18, 19]. Moreover, 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 instructed to design agents that play the resource allocation game. Each student was responsible for designing her own agent. Each group received a different version of the resource allocation game. Each version was governed by different dependencies between the players (that is, 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 her 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.

The line of research most similar to this work includes studies that explore human strategies using a strategy method[3], where people are able to revise their strategies based on their performance in the past [18, 23]. 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. Despite the similarity, asking subjects to design and program agents is different from the strategy method. Developing agents enables subjects to implement much more complex strategies (e.g., using heuristics, learning algorithms, etc.). Moreover the strategy method is usually used to elicit subjects' strategies in games that are relatively simple (e.g. public goods, generations games [14, 18]). Our work, on the other hand, involves a richer and more complex strategy space.

Our results show that the subjects fundamentally modified their agents' strategic behavior throughout the tournaments. Surprisingly, though, 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 in this game (see Sect. 3.2 for an analysis of the equilibrium strategy). Thus, the decrease in the agents' average scores warrant an explanation. One explanation can be given by analyzing the agents' strategies. More agents were less cooperative and their reliability level decreased (i.e., agents did not live up to their promises) throughout the tournaments.

Despite the decrease in the agents' reliability throughout the tournaments, even in the final tournament there were still subjects whose agents were reliable. Among these subjects and others, we observed an interesting behavior pattern: subjects learned from their agents' experiments in the previous tournaments and increased their agents' means of protection against deceiving agents (that is, agents that promise a transfer of resources but do not carry it out). For example, some subjects developed agents that do not send any resources to an agent whose behavior in the negotiation is suspicious (e.g., agents that are too generous and offer to send all their extra resources in exchange for the resources they need).

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 a substantial effect on the market and on agent design.

The rest of this paper is organized as follows. We begin by reviewing related work in Sect. 2. Then we continue with a description of our trading framework in Sect. 3. The experiment design is provided in Sect. 4 and the results in Sect.

---

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

5. In Sect. 6 we discuss our results, and finally we conclude and discuss future work in Sect. 7.

2 Related Work

In this paper we explore the strategic behavior of Peer Designed Agents. A very known competition which has shed light on automated trading agents' strategies is the Trading Agent Competition (TAC) [16, 25, 26]. In this competition entrants develop travel agents that need to arrange itineraries for a group of clients who want to fly from one city to another within a certain time period. Most of the papers describe the strategy of a single TAC agent (e.g., [4, 12]). However, some of these papers describe the strategies of trading agents that were developed for a certain competition. In addition, 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.

Another known tournament is the Prisoner's Dilemma [19]. In the Prisoner's Dilemma the "cooperation" option can be compared to the decision whether to send the promised resources in our domain and similarly the "defect" option is comparative to choosing not to send. In the finite repeated Prisoner's Dilemma problem the equilibrium strategy is to always defect. Likewise, in our setting, if the number of interactions between the players is finite, the equilibrium is that no exchange will occur. However, it has been shown in the Prisoner's Dilemma that if the game is repeated many times cooperation between players can transpire [2, 3]. Actually the winner's strategy in Axelrod's tournament was the *Tit-for-Tat* strategy. According to this strategy, the player cooperates on the first move, and then does whatever her opponent does. In our setting this is equivalent to sending the promised resources in the first agreement and then imitating the other player (that is, sending the promised resource according to the agreement, if the other player fulfills her previous agreement). Nevertheless, in our work we consider situations, where the number of resources that each player needs from the other player is relatively small and consequently the number of iterations is relatively very small. Accordingly, the Tit-for-Tat strategy is not necessarily recommended for our setting, as the number of interactions between players is too small in order for trust to emerge.

Some research in economics has explored the strategic behavior when people are able to revise their strategies based on their performance in the past [24]. Most of them used a strategy method[4]. The strategy method was first proposed by Selten [22]. Similar to developing agents, using the strategy method requires

---

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

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. Strategy method has been applied in many experimental settings. In addition, in some of these works, similar to ours, subjects have the opportunity to revise their strategies based on the feedback received [15, 23]. For example, Selten *et al.* [23] investigate subjects' strategies for playing any 2-person 3x3 games with integer payoffs between 0 and 99. The authors examine the subject's strategic behavior throughout five tournaments. They found that the frequency of the play of pure strategy equilibrium increases throughout the tournaments. Keser and Gardner [15] examine the strategic behavior of experienced subjects throughout three tournaments in a common pool resource game. The authors' main result showed that, even though at the aggregate level the sub-game perfect equilibrium organizes the data fairly well, fewer than 5% of the subjects behave according to the game equilibrium prediction. Despite the similarity, asking subjects to design and program agents is different from the strategy method. Developing agents enables subjects to implement much more complex strategies like using heuristics, learning algorithms, etc. Moreover 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.

The resource allocation game is a simple version of the CT game that was designed by Kraus and Grosz to investigate properties of decision making strategies in multi agent situations [10]. The authors compare the behavior of subjects that play the game online with Peer Designed Agents' that were developed by the same population. The authors show that people design agents that are less cooperative and their strategies are closer to the equilibrium strategies. In contrast to our work, in the CT game subjects design their agents only once and do not have the opportunity to modify them. We on the other hand investigate agents' strategic behavior over time. By doing so our model resembles real life situations, where people can see how their agents perform and modify them according to their performance.

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 [17].

We begin by describing the environment and we continue by analyzing 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 send $Ogive_j^l$ back to her. Player j should inform player i whether she accepts or rejects the offer. Next, 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, i.e. 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 her promise. The performance of each player is determined by her score at the end of the game. The score of a player takes into account both the number of resources she possesses, as well as whether or not she has reached the goal. For each resource the player possesses at the end of the game, she will receive a score of $Score_{Res}$. In addition, if she holds her whole target set, she 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 her to reach the goal, the other player lacks two specific resources, and she 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 she needs to complete the target set, and the independent player can increase the number of resources she 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 her 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 her 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, each possesses books that the other needs for her research, while each of them needs a different amount of books from the other. In this case both can gain from the exchange.

In the following subsection we analyze the different strategies in both configurations.

3.2   Strategies Analysis

The agent's strategy includes offers, responses and decisions regarding which resource to send. These decisions are not separate from one another. Consider an agent that does not send any resources. In such a case this agent can proffer attractive offers (e.g. give all its extra resources in return for the resources it needs to complete its target set) as it does not send the promised resources.

Using a backward induction we can deduce that the equilibrium strategy of an agent (in both roles and configurations) is to send nothing. In fact any agent that does not send any resources is an equilibrium agent, regardless of her negotiation policy. Indeed, not sending any resources is an equilibrium strategy but it is not a dominant strategy (like the "always defect strategy" is not a dominant strategy in the finitely repeated Prisoner's Dilemma [7]). To explain this, consider the OIP configuration and assume agent i plays against agent j. Assume that agent j's strategy is to negotiate one resource at a time. In addition this agent will fulfill its agreement until it has been deceived. Once agent j has been deceived, it stops sending resources. This implies that if agent i is an equilibrium agent, agent i will never complete its target set (as agent j will send agent i only one resource; after sending this resource it will stop sending resources as it has been deceived by agent i). Nevertheless if player i is not an equilibrium agent and sends the promised resources according to the agreement it will complete its target set (as agent j will fulfill its agreement with reference to the second resource).

Similar to the finitely repeated Prisoner's Dilemma, the resource allocation game has a dominant strategy in the final round of the game. In this round the dominant strategy of both players is not to send any resources, since breaking an agreement in this round can only increase the player's score.

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 each group received a different configuration, as described earlier. 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 anonymous 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 strategies.

After a given tournament each of the subjects received feedback about her agent's performance during the tournament. More specifically, each of the subjects received a log file for each game that her 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; thus 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 would be 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 the subjects to put forth effort also in the later tournaments, we explained that their grade would be calculated according to the score obtained by their agents accumulated over all tournaments, whereby later tournaments would receive much larger weights. Moreover, in order to ensure that in each tournament each of the subjects would do their best and would 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 final one.

5 Experimental Results

The agents participated in four tournaments, enumerated 1-4. To analyze the performance of the agents we used two benchmark scores. The first benchmark was 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_{2Res} and sc_{1Res}, 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 was 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 began with at the start 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, and thus $sc_{dep} = 560 - sc_{ind}$. This implies that the average score, denoted sc_{OIPAvg}, 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 both agents complete their target set at the end of the game, it follows that $sc_{2Res} + sc_{1Res} = 560$ and the average score, denoted sc_{ADAvg}, 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. Table 1 summarizes these benchmark agents' scores.

| Agent		Equilibrium	Pareto-optimal
OIP Configuration	sc_{ind}	280	$280 \leq sc_{ind} \leq 330$
	sc_{dep}	80	$560 - sc_{ind}$
	sc_{OIPAvg}	80	140
AD Configuration	sc_{2Res}	80	$230 \leq sc_{2Res} \leq 330$
	sc_{1Res}	80	$560 - sc_{2Res}$
	sc_{ADAvg}	40	140

Table 1. Benchmark agents' score.

First we will examine the change in the agents' average score for all the tournaments, as depicted in Figures 1(a) and (b). When comparing the scores to the benchmarks we can see that on average the PDA's obtained 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.
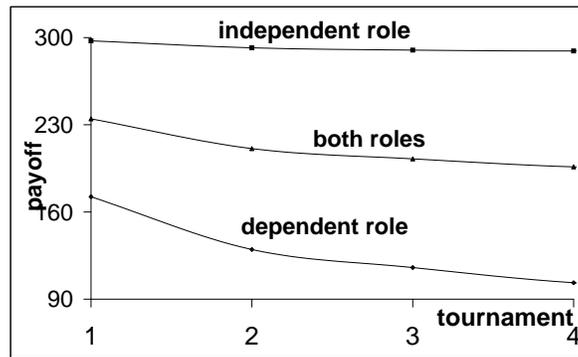
(a)
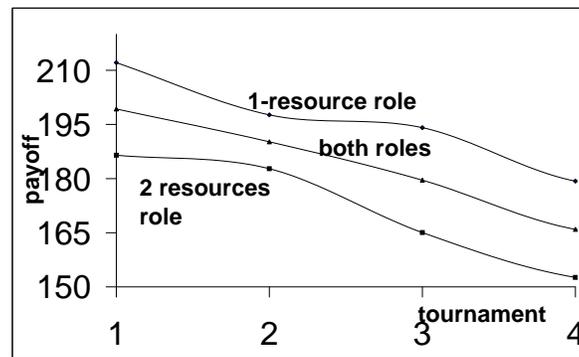
(b)

Fig. 1. The agents' average score in the different tournaments (a) in the OIP configuration (b) in the AD configuration.

Figure 2 presents the percentage of agents (not including agents playing the independent role) that completed their target set. We can see that in all consid-

ered roles the percentage of agents that completed their target set monotonously decreases over the tournaments. We observe 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 completed their target set can be 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, $p < 0.01$).
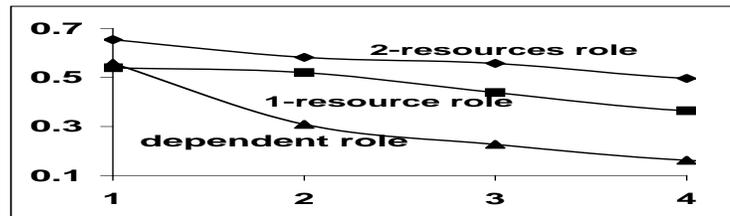

Fig. 2. The percentage of agents that complete their target set.

Two possible reasons can explain the decrease in the scores and the percentage of agents that reach 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 it. 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). We discuss this strategy in detail in section 6.1. The second reason is that agents tended to break more agree-

ments 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. We elaborate on this behavior in Sect. 6.2.

6   Discussion

In this section we present a list of the possible reasons for our results based on an analysis of the subjects' codes.

6.1   Agents' Strategies for Sending Resources

It is interesting to unveil the reasons for the decrease in the agents' average scores. We investigated the agents' strategies for sending resources (i.e., deciding whether to send resources and if so how many) by reviewing the agents' codes. We observed that about 50% of the subjects decreased their agents' reliability throughout the tournaments in both configurations. An additional observation can be made if we also divide the agents into two types: (a) agents that never send the other agent all the required resources it needs to complete its target set (note that a special subset of these agents are equilibrium agents), denoted NSRR (Never Send Required Resources), and (b) the complementary set of NSRR, denoted cNSSR. That is, agents for which there are scenarios where they send all the resources that the other agent needs to complete its target set. As Table 2 shows the number of NSSR agents increases slightly throughout the tournaments in the AD configuration, and it increases more significantly in the OIP configuration. Although the agents' strategies in the last tournament are closer to the equilibrium strategies than in the first tournament, most of the agents in the last tournament are not equilibrium agents.

Our findings indicate that subjects develop agents that deceive more and break more agreements throughout the tournaments although their strategies even in the last tournament are still different from the equilibrium strategies. As a matter of fact in both groups about 33% of the agents usually do not deceive without being deceived first (some of the subjects even explicitly wrote that they do not wish to deceive).

6.2   Detection and Protection Against Dishonest Agents

We also observed that subjects increased their means of protection against deceivers. When we analyzed the modifications the subjects had implemented in their code, we were able to classify this type of behavior into four sub-categories:

1. Punishing the other agents based on their behavior in previous rounds (e.g., the agents cease to send any resources to agents that have priorly deceived them).

	AD configuration	
Agent's type	first tournament	last tournament
NSSR (EQ)	5 (0)	6 (1)
cNSSR	12	11

	OIP Configuration	
Agent's type	first tournament	last tournament
NSSR (EQ)	3 (2)	9 (6)
cNSSR	12	6

Table 2. The change in agents' types throughout the tournaments.

2. Limiting the number of resources which they agree to negotiate per round. Using this protection method allows the agent to minimize the risk of playing against deceivers by avoiding sending all the resources required while not receiving any resources in return.
3. Postponing sending the resources the agent has committed to send during round r to the following round $r + 1$, after being assured it has acquired the promised resources from the other agent.
4. Trying to characterize the other player as a deceiver based on its suspicious negotiation policy (e.g., the agents will not send resources to agents that are too generous and offer to send all their extra resources in exchange for the resources they need).

Subjects were able to make these modifications thanks to the large and rich game strategy space that our game introduces. In particular, 33% and 64% of the subjects in the OIP game and the AD game, respectively, improved their agents' protection level using the means of protections detailed above.

6.3 Does Modifying the Agent Ensure Improvement?

The final issue we investigated is whether the modifications made by the subjects in their agents actually yielded agents that performed better. In other words for each agent A_j we checked which agent version, after tournament T_i, would receive a higher score, the older version of the agent, which was submitted for tournament T_i, or the newer version, which was submitted for tournament T_{i+1}, when competing against the same agents that participated in tournament T_i.

Despite the fact that the average agents' score decreases over tournaments (as illustrated in Figure 1) this does not necessarily imply that subjects made their agents worse. This is because the new agent's version can increase its score at the expense of the other older version of agents by being less cooperative. As a result the new agent (the agent that was sent to tournament T_{i+1}) will achieve a better score when it plays against the old version of agents (all the

other agents that were sent to tournament T_i). However, when the new version agents compete against each other, since they are less cooperative, the total score of all agents decreases. This phenomenon is similar to the Prisoner's Dilemma [19]. Assume that all the agents' original strategy in the Prisoner's Dilemma is always to cooperate. Given this, each agent that will modify its strategy to always defect will increase its score (since the other agents' strategies are always to cooperate, when this agent plays against all the other agents it will achieve a better score). However, the average score of an agent in the new tournament, which includes agents whose strategy is to always defect, is much lower than the average score of an agent in the old tournament, which includes agents whose strategy is to always cooperate.

Figures 3(a) and 3(b) show that a higher proportion of subjects improved the performance of their agents. Thus, we can infer that the decrease in an agent's average score over tournaments cannot be attributed to the hypothesis that humans made their agents perform worse. Instead, and based on the results, we can attribute the decrease in the agents' average scores to the decrease in their reliability.

7   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, 5, 11, 18, 19]. 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 each player's dependency on the other player.

Even though the basic configurations are different, the results of our experiment show that the agents strategic behavior was fundamentally modified throughout the 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 during the course of the tournaments and as the tournaments progress they are less likely 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.

References

1. J. Andreoni. Why free-ride? strategies and learning in public goods experiments. *Journal of Public Economics*, 37:291–304, 1988.
2. R. Axelrod. *The evolution of cooperation*. Basic Books, 1984.
3. R. Axelrod and W. Hamilton. The evolution of cooperation. *Science*, 211:1390–96, 1981.
4. M. Benisch, J. Andrews, and N. Sadeh. Pricing for customers with probabilistic valuations as a continuous knapsack problem. In *Eighth International Conference on Electronic Commerce*, pages 128–134, 2006.
5. 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.
6. C. F. Camerer and E. Fehr. *Measuring Social Norms and Preferences using Experimental Games: A Guide for Social Scientists. Foundations of human sociality: Economic experiments and ethnographic evidence from fifteen small-scale societies.* Oxford University Press, New York, 2004.
7. K. Chellapilla and D. Fogel. Evolution, neural networks, games, and intelligence. *Proc. IEEE, Vol.*, 87(9):1471–1496, 1999.
8. C. Da-Jun and X. Liang-Xian. A negotiation model of incomplete information under time constraints. In *AAMAS02*, pages 128–134, 2002.
9. I. Erev and A. Roth. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibrium. *American Economic Review*, 88(4):848–881, 1998.
10. 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.
11. 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.
12. M. He, A. Rogers, X. Luo, and N. R. Jennings. Designing a successful trading agent for supply chain management. In *AAMAS06*, pages 1159–1166, 2006.
13. Y. Ishihara, R. Huang, and J. Ma. A real trading model based price negotiation agents. In *AINA*, pages 597–604, 2006.
14. C. Keser. Strategically planned behavior in public goods experiments. Working Paper, Scientific Series 2000s-35, CIRANO, Montreal, Canada, 2000.
15. C. Keser and R. Gardner. Strategic behavior of experienced subjects in a common pool resource game. *International Journal of Game Theory*, 28:241–252, 1999.
16. P. L. Lanzi and A. Strada. Analysis of the trading agent competition 2001. *SIGecom*, 3(2):1–8, 2002.
17. E. Manisterski, R. Katz, and S. Kraus. Providing a recommended trading agent to a population: a novel approach. In *IJCAI 2007*, pages 1408–1414, 2007.
18. T. Offerman, J. Potters, and H. A. A. Verbon. Cooperation in an overlapping generations experiment. *Games and Economic Behavior*, 36(2):264–275, 2001.

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

20. A. Roth. Bargaining Experiments. In *The Handbook of Experimental Economics, H. Kagel and A. E. Roth (eds.)*. Princeton University Press, 1995.

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

22. R. Selten. Die strategiemethode zur erforschung des eingeschrnkt rationalen verhaltens im rahmen eines oligopolexperiments. *Beitrge zur experimentellen Wirtschaftsforschung*, 1:136–168, 1967.

23. R. Selten, K. Abbink, J. Buchta, and A. Sadrieh. How to play (3x3)-games: A strategy method experiment. *Games and Economic Behavior*, 45(1):19–37, 2003.

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

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

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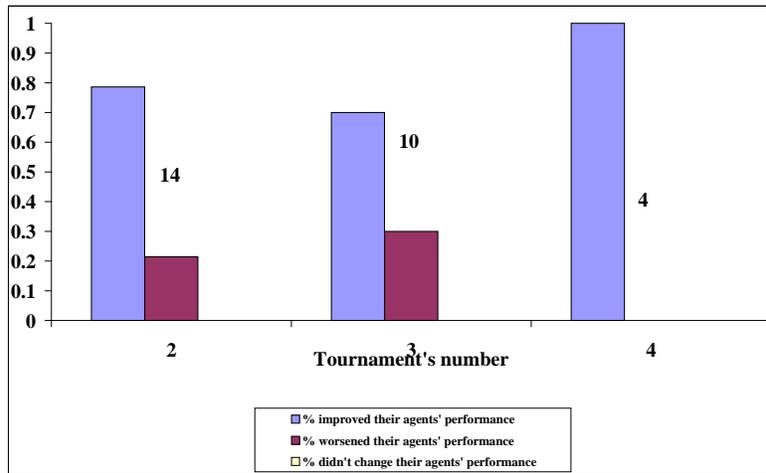
(a)
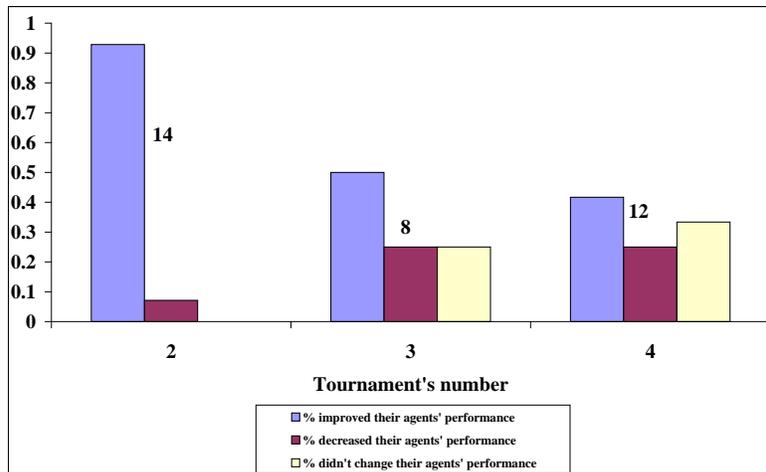
(b)

Fig. 3. Changes in agents' performance in (a) the OIP configuration (b) the AD configuration among modified agents with respect to the original agents. The numbers near the bar represent the number of subjects who modified their agents.