Ben-Gurion University of the Negev

Faculty of Engineering Science

Department of Information Systems Engineering

Coordination and Collusion in 3-Player Strategic Environments

Thesis submitted in partial fulfillment of the requirement for the degree of master of science in the Faculty of Engineering

Submitted by

Yaniv Mazliah

Department of Information Systems Engineering

Ben-Gurion University of the Negev

E-Mail: yanivmaz@post.bgu.ac.il

Supervisor

Dr. Ya'akov (Kobi) Gal

Department of Information Systems Engineering

Ben-Gurion University of the Negev

E-mail: kobig@bgu.ac.il

Contents

Abstract

Coordination is the act of making participants in an activity agree to work together in a way that benefits the group. There is inconclusive evidence as to whether having people train by practicing tasks with computer agents can help them improve their coordination. In this thesis we empirically investigated this question by means of extensive experiments involving human subjects who trained with computer agents to play a three-player coordination task that is a common test bed in AI to evaluate computational cooperation strategies. Two groups trained to play the game each by means of a different method. The first group trained by playing the game with other people and computer agents. The second group trained by designing agents that would participate in a tournament against each other. Following the training, we compared the performance of subjects playing state-of-the-art agents from the literature. The results revealed that training with computer agents (or designing such agents) improved people's performance compared to training by playing against state-of-the-art agents. In addition, a cognitive hierarchy analysis of people's behavior revelaed that people who trined using computer agents were more likely to exhibit sophisticated reasoning than those who trained with people. These results demonstrate the efficacy of using computer agents as tools for improving people's skills when interacting in strategic settings. This can save considerable effort and can result in better performance than people training with human counterparts.

Chapter 1

Introduction

Environments in which people and computers cooperate emerge in a wide variety of application domains (e.g., hospital care-delivery systems, systems administration applications) as well as in virtual reality and simulation systems (e.g, disaster relief, military training) [13, 21, 23, 26]. Automated computer agents in these settings are designed to supporting people, and they acting as proxies for individuals or organizations, or work autonomously to carry out the actions for which they are responsible.

A breadth of prior work exists on the subject of designing and evaluating computer agents who cooperate with people. However the effects of using computer agents to change people's behavior in strategic settings is inconclusive. Autonomous agents designed by researchers and students commonly use opponent modeling, game theoretic reasoning and machine learning approaches that allow them to perform successfully in their respective settings [19]. However when deciding whether to cooperate, people prefer to cooperate and coordinate with other people rather than with computer agents [22].

To address this gap, we investigated the question of whether people training with automated agents can improve their performance in a representative setting involving coordination among groups of three participants. Our method involves training people to coordinate in a given task with other participants, whether people or computer agents. We compared people's behavior during training with that of their performance during a separate testing phase conducted on the same task. The challenge in evaluating people's performance in these multi-participant task settings is that their behavior depends in part on the strategies of the other participants. We therefore used a standardized agent to interact with people when comparing their performance in the testing phase. This agent was chosen from the state-of-the-art, namely its proficiency was already demonstrated when interacting with other computer agents (or people) in other studies. The use of a standardized agent provided an objective metric with which to evaluate people's performance.

Our empirical methodology consisted of a three-player multi-round coordination game of imperfect information that is used as a research test-bed for evaluating novel agent designs in Artificial Intelligence, named the Lemonade Stand game [37]. We compared people's performance in these settings when interacting with other people or with another computer agent. All situational contexts were kept identical for subjects in both conditions. Thus, any difference in their behavior can be attributed to the history of their prior interaction in the training phase. A separate group of students were given a task to design an automated agent that will participate in a contest. These subjects later participated in a test session in order to derive whether this form of training was beneficial or not.

The results show that using automated agents to train people improved their performance in comparison to the standardized agent's performance. Also, training with people improved the performance of those people who coordinated more often with the standardized agent. Further analysis revealed that the top-scoring people preferred to cooperate with the computer agent rather than with other people, causing the average human score to drop. In both settings, there was no significant improvement in people's performance after training with people. Training by designing agents improved people's performance scorewise similar to people who trained with automated agents. However we found that people who trained with agents performed more similarly to those agents than people who designed their own agents.

These results provide insight to agent designers for human-computer decision-making as well as social scientists. The findings suggest that in settings requiring coordination and agreements, people can learn to be more skillful by learning from computer agents. These agents can be used as tools for training people in such tasks. This can result in considerable savings in cost and effort in comparison to training with people.

1.1 Contributions

This thesis makes the following contributions:

- We analysed human behavior in the Lemonade Stand game domain using a cognitive hierarchy model.
- We show that training with computer agents or designing computer agents improves people's performance.
- The findings of this work were published in a journal publication [17] and a workshop publication [20].

Chapter 2

Related Work

This thesis relates to work in several areas: human-computer negotiation, opponent modeling and using computer agents to train people.

2.1 Negotiation

We encounter negotiation situations almost every day of our lives. Haggling over the prices might be the first example which comes to mind, but also deciding which movie to see is a certain type of negotiation. The process of negotiation usually includes the balance between increasing one's benefits from the outcome while still reaching a conclusion. Difficulties in negotiation can result from lacking information (not knowing the other party's interests) or mistrust (not knowing if the other party will honor the agreement). These difficulties increase when negotiation involves people and computers. Past experiments have revealed that people tend to offer less favorable offers to computers [31]. There is also a difference between a one shot negotiation scenario where we have little information about the other party, to repeated negotiation scenarios where we can learn about the opponent but also try to build up our own reputation. Negotiation is a hard task and therefore training may be needed to improve performance. Negotiation can also include more than two parties. For example, in a conflict between two sides, there is often room for a mediator. The mediator's role help the parties in the conflict reach a resolution and improve the outcome for both sides. Previous research [18] has shown that it is possible to use automated agents as mediators between two people in order to improve the outcome of the conflict.

2.1.1 Negotiation between people and automated agents

Past research on negotiation between people and automated agents mainly focuses on the ability of agents to successfully reach "good" agreements in such negotiations. Negotiations with people lacks some basic assumptions that are usually used in research, such as rationality or complete information. People do not always reach an equilibrium solution [9] which could prove problematic to equilibrium based automated agents.

One way of creating an agent that could interact with people is by using a learning mechanism. The "learning agent" is an agent that uses past encounters with people to model their behavior. The model can use different statistical approaches to decide which response is best for a given situation. Past research on negotiating agents shows mixed results. The diplomat [16] agent for example was able to achieve good results while competing with people. It attempts to estimate the personality of its rival and decides whether he will keep his promises or not. Gal et al. [10] shows that it is possible to use a model to learn the preferences of the opponents, and achieve better results than heuristic and rule based automated agents and human players. On the other hand the autONA [2] agent (who uses belief functions based on past negotiations) displayed poor results when trying to negotiate with people; even after the agent was modified it still displayed worse results compared to the human players. More information about negotiation between people and automated agents can be found in the survey conducted by Lin and Kraus [19].

2.1.2 Opponent Modeling

To be able to cooperate with or to counteract opponents we need to be able to anticipate their behavior. For this purpose we can try to model their view of the world. A good way to describe an agent's view of the world is by means of a network of influence diagrams (NID) [12]. The NID graphically describes the relationship between different parameters in the agent's model of the world. Since an agent's decision depends on other agents' actions, we assume some agents build models for other players to be able to predict their actions.

A model of an opponent should be dynamic. Specifically, if we assume other agents are trying to anticipate our action then their strategy should change over time, as they learn and respond to our and other agent's actions. The prediction of the other agent's actions is important in the case where no communication method exists. Even if such communication methods exist, an agent cannot always trust other agents and sometimes there is not enough time to convey the information from one agent to another. Previous research [11] has shown that by using this method it is possible for automated agents to learn how social factors influence different people, and thereby enabling agents to improve their performance. These agents were also able to outperform agents using traditional game theory strategies.

2.2 Training people with computer agents

The use of automated agents as trainers is very appealing. First unlike people an automated agent is always available. Second an automated agent can be used in several sessions to train multiple subjects. And lastly after it has been designed the cost incurred for its deployment is very low. One important aspect of a training agent is its balance between forgiveness and punishment for mistakes. Punishing for making mistakes encourages players to improve, however without some forgiveness players might become frustrated and quit their training. An ideal training agent should lead a trainee to situations where he can learn the most [8]. Another challenge in using automated agents is that it is not always possible to create an agent that outperforms people. For example, in chess though there are programs that can beat people [4], in the game of “Go” automated agents still have difficulty winning against people (even when handicapped) [24]. Evidence shows that training is beneficial to an individual’s performance, motivation and sometimes causes empowerment [1]. People who train gain confidence in their abilities and also gain experience that can serve them later outside the training environment. To maximize the benefits of training it is important to motivate the trainee to be willing to train. It seems that motivation to train has a major influence on the trainee’s success in training. To increase the participants’ motivation it is important to create an environment suitable for training, reduce the participant’s anxiety and set goals to the exercise [5] [15].

One setting in which agents have been used to train people is virtual environments. Virtual environments are a means of simulating real world behavior using computer resources. These environments are used for various applications, such as video games, working tools (word processors and virtual desktops) and real world simulation tools. One particular use of virtual environments is training people, by using simulation to create a close clone to real life situations. Training by means of virtual environments improves people’s perception of certain situations by giving them experience that they could not otherwise have gained. For example it is possible to efficiently train fire-fighters [29] using virtual environments without risking lives or property. Other examples include virtual environments used for medical research [33], or environments with unilateral spatial neglect for training to cross streets in a safe manner.

In some environments several agents are active and influence one another. This is common in strategic environments, thus additional players are needed to create a simulation. We hypothesize that playing with other players could prove an invaluable experience; having the opportunity to feel the environment first hand could help people gain insight and help them develop a better approach to the environment. Since we are discussing virtual environments it is possible to use either virtual or real players. The advantage of the virtual agents is that they are more controlled; nonetheless, they need to be designed and adapted to the environment. People are more flexible but usually are harder to entice and will need an incentive to participate in a simulation. Previous research shows that the use of virtual environments is transferable to real world situations [32] [25], and can also be used as an educational tool [6, 27]. In this thesis we used the Lemonade Stand game domain which is a virtual environment used as a test bed in AI to evaluate computational cooperation strategies.

Surprisingly very little work has been done to measure the effect of simulation with automated agents on people's performance. Some work has been done in the field of negotiation. Druckman et al. [7], for example, evaluated the influence of simulation in training students in the skills of negotiation. In their study, they found that designing a simulation was more effective than participating in one. Kenny et al. [14] and Traum et al. [30] used virtual humans to facilitate people's negotiation, leadership and interviewing skills. These virtual humans were tested in several negotiation scenarios in social and military contexts. We hypothesize that participating in a simulation with automated agents or designing agents for such simulations can improve people's performance, and as such is a good way to train people.

Chapter 3

The Lemonade Stand Game Domain

In this section we describe the lemonade stand game which served as the domain of our study. The Lemonade Stand game [36] describes a three player game, where each player is a lemonade vendor who tries to maximize his profits. The rules of the game are:

- In the Lemonade Stand game there are n beaches (usually n=12) , where each beach is a possible place for a player to place his lemonade stand.
- The island is round and the beaches can be seen as though they are on a clock, as depicted in the example given in Figure 3.1.
- Each round the players choose their locations on the board simultaneously.
- The players in the game represent lemonade vendors who compete to serve customers in their vicinity.
- The people on the island will buy lemonade from the closest stand.
- Each player's score is the distance to his most right neighbor plus the distance to his most left neighbor.
- This is analogous to the profit each player would make if the lemonade customers are uniformly distributed around the island.

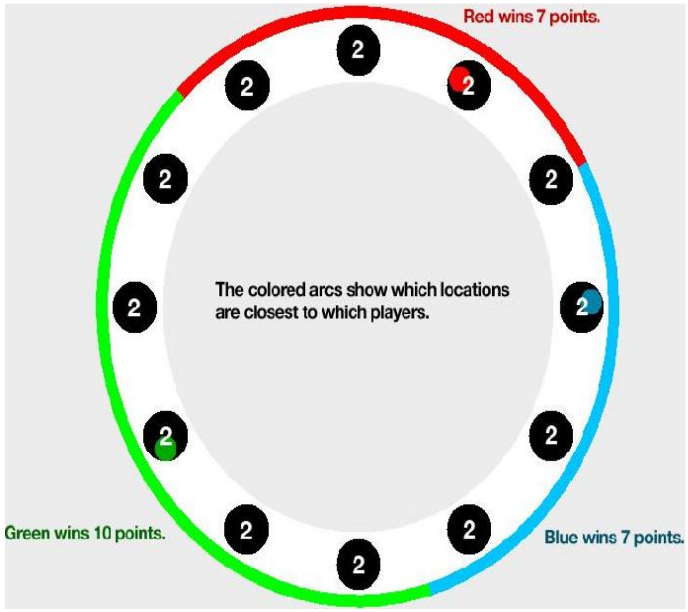


Figure 3.1: The browser version of the Lemonade Stand game used in this research.

- If there are two players at the same spot their score will be 6 and the remaining player will score 12.
- If all of the players are at the same spot each will score 8 points.
- The game is played over several rounds usually 100 rounds.

The game features competition between players where each player's profit is deduced from another player's profit. Another aspect of the game is that it allows players to cooperate. For example, two players can set their stands across from each other to maximize their score at the expense of the third player.

3.0.1 Strategies in the game

In this section we describe common strategies for playing the game [34]. A visualization of these strategies on a Lemonade Stand board game is shown in Figure 3.2. In one of these tactics, called "Stick", a player chooses to remain in the same location for two or more

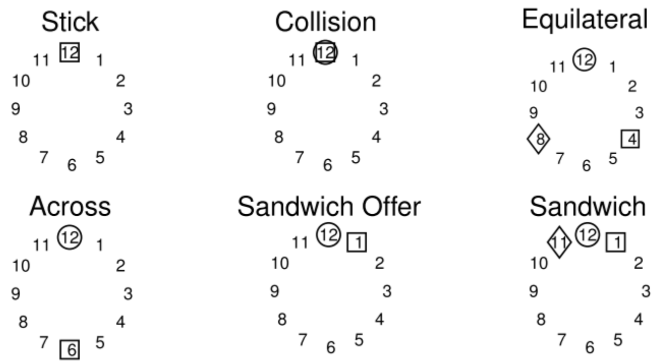


Figure 3.2: Key strategies in the Lemonade Stand Game.

consecutive rounds. Another tactic, named “Follow”, is a strategy where a player chooses a location that is directly across from the location of another player in the previous round.

We will now explain several terms and strategies in the Lemonade Stand game. The situation in which both players are positioned directly across from one another is called “Across” (seen in figure 3.2). This strategy guarantees a payoff of between 6 and 12 to each of the players that are situated across from each other, while the third player, who is “locked in”, receives a payoff of 6. The “Equilateral” positioning (seen in figure 3.2) in which the players are situated at an equal distance from each other results in a score of 8 points for each player. Another more aggressive form of cooperation is the “Sandwich” strategy(seen in figure 3.2), in which one of the players is positioned directly between the two other players. This strategy provides a higher outcome of 11 for both of the agents positioned at the edges and a low score of 2 for the agent in the middle. To achieve a sandwich outcome in the game a player might try to use a “Sandwich offer” (depicted in Figure 3.2).If a player identifies a “sticking” player he can move to the location next to him, and effectively offer the third player the option to perform a “Sandwich” collaboration move. “Collision” (illustrated in Figure 3.2) is a scenario where two players choose the same position. While this move forces both players to receive a low score of 6, it could be used to encourage a player to leave a certain position.

3.0.2 Nash equilibrium in the Lemonade Stand game

In game theory a Nash equilibrium is a solution for a game where each player has nothing to gain by changing his own strategy, as long as the other players will not change their own. There are several classes of Nash equilibria strategies in the Lemonade Stand game. One of these classes includes situations in which three players are located in different locations on the board and each player earns at least 6 points. This is because no player can increase its score by changing his position. For example, the equilateral positioning is a Nash equilibrium of this class (changing the position of any one player will not increase its score). This might seem like a good strategy profile for the game. However, the across strategy profile is also a Nash equilibrium, and dominates the equilateral positioning from the point of view of the two players situated across from each other. Each of the players would earn 9 points on average instead of 8 points in the "equilateral" Nash equilibrium. Another class of Nash equilibrium strategies includes situations in which two players are located in the same position, while the third is positioned directly across from them. The "Sandwich" strategy profile is not a Nash equilibrium, because the low scoring player in the middle has an incentive to deviate from its position to increase its score(in Figure 3.2 the "circle" player in the sandwich can move to position 6, for example, to increase its score from 2 points to 10 points).

3.0.3 The annual Lemonade Stand game competition

The first step in this domain was to see if computer agents can coordinate with each other. This was done by arranging a Lemonade Stand game tournament, where agents were developed independently [37]. The motivation was to test coordination between agents in an AI environment. Two types of cooperation can be defined in the Lemonade stand game: stable cooperation and unstable cooperation. Stable cooperation is when two players collaborate to increase their score but the game remains in a Nash equilibrium state. In the Lemonade Stand game stable cooperation would be to play from across each other leaving the third agent with the score of 6. Unstable cooperation would be to try to put the game in a nonequilibrium state to gain a higher score.

The results of the tournament showed that in most of the situations unstable situations did not occur (less than 2%). This seems to indicate that either such opportunities do not exist, or that the players decided that it was not worth pursuing such opportunities (i.e. The reward was not worth the investment). The analysis of a Lemonade Stand competition using automated agents [35] shows that agents who focused too much on learning and that were not quick enough to respond to other players, were considered less cooperative by the competitors and thus suffered lower scores.

The Lemonade Stand game tournaments are still conducted yearly. To keep the game interesting each tournament is different than the one before it. For example the 2012 tournament allowed agents to keep information between games, enabling agents to learn which strategies are present in the tournament. And in the 2011 tournament, the settings of the island were changed by dividing the customers in a nonuniform way on the different beaches of the island.

3.0.4 Lemonade Stand agents in the literature

In this research we used several agents that previously participated in the annual Lemonade Stand game tournament. We will now explain the role of the agents in this research and their operating method in the Lemonade Stand game.

EA^2 the winning agent in the 2010 tournament

Since it was necessary to have automated agents in our domain (we needed a standard agent to compare with people's performance), we conducted a local tournament using all of the available agents that ever participated in the Lemonade Stand tournament. Official tournaments between agents run thousands of rounds. However since we experimented with people we ran shorter sessions of 30 rounds. After testing we verified that the EA^2 [28] still had the best performance even with shorter sessions. The EA^2 strategy consists of trying to cooperate with one opponent by sitting across from its position. The EA^2 tries to define for each opponent whether it is a “Follow” (moves to a position across from an opponent) type or a “Stick” (remains in the same position) type. The EA^2 then decides which of its two opponents will more consistently to be one of these types and cooperate accordingly.

The training agents

After we selected a standard agent we chose two automated agents to train people. The selected agents were the runner ups from the tournament we conducted, i.e., “Goffbot” and “MatchMate”. These agents took part in a previous annual Lemonade Stand game tournament [37]. We decided to use different agents to train people than the one we used in the testing session. We hypothesized that using the same agent in both training and testing could lead to people to learn to play better against a specific opponent. The “Goffbot” agent used an “Across” strategy where it tried to cooperate with the opponent with the lowest score. It also tried to identify “Sandwich” attacks made on itself and attempted to avoid them. The “MatchMate” agent tried to play the “Follow” action as fast as possible in order to cooperate early.

Chapter 4

Empirical Methodology and Results

This chapter will explain the empirical methodology we used in this research. Our study had two phases of experiments that tested people's behavior and performance in the Lemonade Stand game. In the first phase we allowed people to play a single session of the game. In the second phase, in order to improve their performance, we trained people by means of automated agents and by allowing them to design automated agents. We than compared their performance using the standardized agent EA^2.

4.1 Empirical methodology - First phase

For the first phase of the experiments we recruited subjects to play a single session of the "Lemonade Stand game" using the Amazon Turk. The subjects were given written instructions on how to play the game as well as a quiz at the end to ensure that each subject understood the rules (The instructions can be found in Appendix A.1). The players were paid according to their performance(2 to 10 USD). Since several subjects provided feedback stating that the game seemed to be too random, we added a brief explanation of basic strategies for the game (see Appendix A.2). The participants then proceeded to play the game. In this phase 31 subjects played a single session with 2 automated agents("EA^2" and

"Goffbot"), and 32 subjects played a single session with a single automated agent("EA^2"), i.e. two subjects played with one agent. The subjects were told that they might play with other people or t with automated agents, during the experiment. Thus during the games they did not know who with whom they were playing. Subjects were paid according to their performance, where each player was paid according to his individual score regardless of how other players scored(25 to 45 NIS, i.e., 5 to 10 USD).

The goal of this phase was to learn how people adapt to the Lemonade Stand game domain. Nonetheless, it was also interesting to see if the automated agents could perform well in an environment with people.

4.1.1 Results: First phase of the single session experiment

The results of the single session experiment showed that people perform poorer than agents in this environment. Since the Lemonade Stand game is a zero sum game (the sum of the rewards to all players is always 24), the average expected results of each player is 240 points.

We first discuss the games with the participation of two automated agents and a single human subject. For this situation the results were one-sided towards the automated agents. The two automated agents were able to quickly find a way to cooperate with the "Across" method, and the human player was left with a very low score. The average score of the automated agents was 264.3 while the average score of the human players was a mere 191.8. The automated agents significantly outperformed people in this environment.

The second environment was with one automated agent and two human subjects. In these games the results were less one sided than in the previous games. Nonetheless, the automated agent still achieved a better score overall than the human subjects. The automated agent had less success at finding a partner with whom to cooperate (in an "Across" fashion). In some of the games in which the subjects formed an "Across" collaboration with the automated agent, they broke it after a few rounds by moving to a different location.

Overall the average score of the automated agent was 251, and the average score of the human players was 234.5. The automated agents significantly outperformed people in this version of the game as well.

We attribute the large gap in the scores to the fact that people were new to this domain with no previous experience in this environment. We analysed subjects' behavior during the game and found that most of the time people tend to change positions. Consequently, because people do not stay in one place it is very hard to predict their next move, making cooperation very difficult. A summary of people's behavior is detailed in Table 4.1. The table shows the average results for human behavior. "For example, when playing two agents people stay in the same position(Stick) 4.8 times on average in a single game. We can see that playing with one agent increases the amount of times people "Stick" and "Follow" wherein people were able to achieve a higher score. This can be explained by the fact that the agent needs to cooperate with an opponent and form a partnership with one of the human players. However, in the two agent scenario the two agents ended up cooperating with each other in almost every game, resulting in low scores for people.

We also noticed a very interesting phenomenon in people's playing strategies wherin people tend to use a very basic heuristic in some of their moves. This heuristic is simply "move to the middle of the large gap between the last position of the two opponents" as depicted in Figure 4.1. This heuristic assume the opponents will most likely stay in the same position or close to that position; it also splits the points of that gap evenly between the other two players. If all players use this heuristic the only stable solution for the game is an equilateral move (where each player earns 8 points). Subjects used the heuristic on average 10 times when playing with one agent and 7.8 times when playing with two agents, as shown in Table 4.1.

The results of the first phase revealed that people achieved a statistically significantly lower score than the automated agents. Our analysis shows that people do not use the stick action enough which makes them unpredictable and hampers cooperation. We noticed that

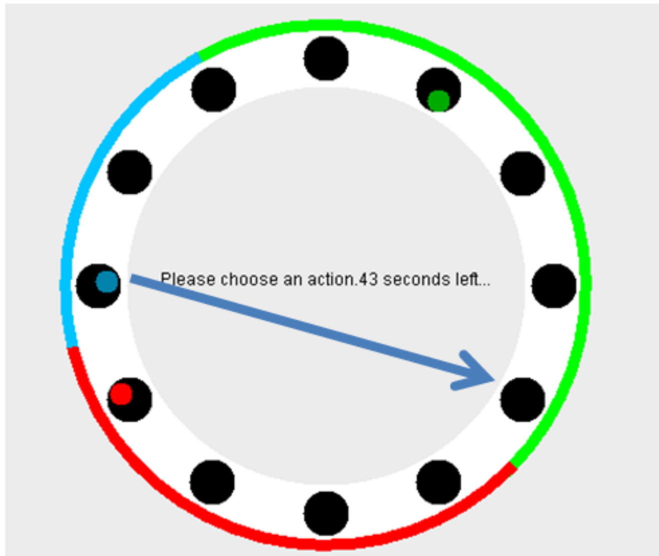


Figure 4.1: A basic Heuristic people use, i.e., move to the middle of the large gap between opponents.

	Score	Stick	Follow	Heuristic
Two Agents	191.8	4.8	6.6	7.8
One Agent	234.5	8	11.7	10

Table 4.1: Summary of human behavior and scores in the single shot experiment

people use a heuristic wherein the strategy does not rely on cooperation (the player tries keep its distance from at least one player). Therefore we decided to try to train people to improve their performance.

4.2 Empirical methodology - Second phase

We recruited 56 undergraduate students from Ben-Gurion University to play the game. Ages ranged from 24 to 30, 58% males and 42% females. All subjects played 90 rounds of the Lemonade Stand game, divided into three epochs of thirty rounds each. The first two epochs (called "training epochs") were used to train the people to play the game, and their performance was measured during the final "testing epoch". The players' configuration in the training epochs varied as follows. In the "all-human" training group, the players' configuration included only human players. In the "single-agent" training group, the players' configuration included two human players and a single agent player. In the "two-agent" training group, the configuration included a single human player and two different and independent agent players. The players' configuration for the testing epoch included two human players and the standardized agent player.

Subjects were randomly divided into the single- and two-agent training groups ("Goff-bot" and "MatchMate" agents), as well as a baseline group in which the subjects played a single testing epoch with the standardized agent("EA^2" agent). Prior to playing the game all subjects were provided game instructions, as well as basic strategies in the game as described in the previous section. Altogether, there were 16 games played by the no-training group, 19 games played by the all-human and single-agent training groups, and 12 games played by the double-agent training group. The subjects were paid according to their total score in all three epochs. Payments were score based and were not related to the scores that other subjects received.

A different group of students were asked to program an agent that would participate in a tournament of the Lemonade Stand game. We hypothesized that programming an agent would help people achieve a better understanding of the elements of the game and thus would help them improve their performance. These subjects did not undergo any training sessions and played a single game with the standardized agent to test their abilities. In the second phase 34 students participated in this type of training. The students partook in this training as a part of an assignment in their course and received bonus points according to their agents' results, in the Lemonade Stand game tournament. Examples of some of these agents' strategies included: staying in the same position unless it had a bad utility, moving across to the player that moves the least, looking for a player who stays in the same position and trying to sandwich it.

4.2.1 Results: Second phase - repeated sessions

All results reported as significant in the following section were confirmed to be within the $p < 0.05$ range by means of single-factor ANOVA tests. We hypothesized that the subjects who trained with two agents would increase their performance in the testing epoch in comparison to subjects who trained (1) with a single agent; and (2) with the standardized agent. We first compared people's games under the various conditions to that of the standardized agent. Figure 4.2 shows the average aggregate performance of people and of the standardized agent in the testing epoch. As depicted in the figure, the EA^2 agent significantly outperformed people in the all-human and all-human training games. However, the difference in performance between the EA^2 agent and people was not significant in the single-agent and double-agent training conditions. This effect was also consistent in the set of games in which people designed their own agent (as illustrated in Figure 4.2 due to overlapping standard error bars).

We conjectured that the reason for the improvement in people's performance was that people learned to play beneficial strategies from interacting with agents in the training epochs. To examine this, we analyzed people's behavior in the game. Table 4.2 shows the

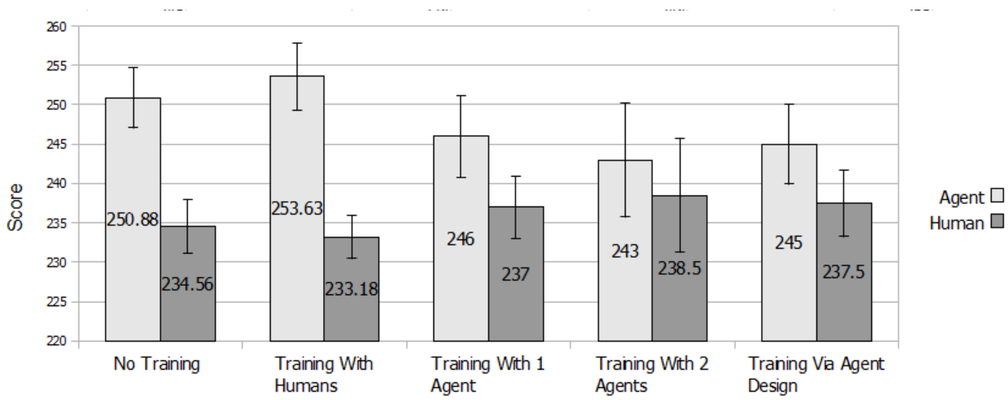


Figure 4.2: Performance comparison: People versus the standardized EA^2 agent

frequency of "Follow" and "Stick" strategies in the testing epoch for the all-human and double-agent training games. As shown in the table, the standardized agent engaged in significantly more "Follow" and "Stick" strategies compared to people in the all-human set of games. Thus, people were less likely to cooperate than the standardized agent during the test epoch, and played more erratically. However, there was no significant difference in the number of "Follow" and "Stick" actions of the standardized agent and people in the double-agent set of games. Consequently, the number of possible across outcomes for the standardized EA^2 agent (also shown in the table) was significantly higher than for people in the no-training condition but not in the two-agent training set of games. As shown by the table, people engaged in significantly more "Follow", "Stick" and "Across" strategies in the two-agent training set of games (shown in boldface) than in the single-agent training set of games. The difference between the number of "Follow" and "Stick" strategies played by people in the single-agent and no-training sets exhibited a similar pattern. Therefore, we attribute the improvement in people's strategies in the single- and double-agent conditions to their increased use of cooperative strategies. Lastly, as shown in Figure 4.2, people were not able to outperform the agent after training. We attribute this finding to the inherent difficulty of playing against state-of-the-art agents in this game.

	All-Human			Two-Agent		
	Follow	Stick	Across	Follow	Stick	Across
People	8.82	7.13	9	**17**	**21**	**25.33**
EA^2	**13.4**	**20.8**	11.58	23.08	21.33	21.92

Table 4.2: Number of "Follow" and "Stick" strategies used by human players in the testing epoch

Comparing people's performance across conditions

We now compare people's performance scores across the various conditions. The number of "Follow", "Stick" and "Across" moves by people in the two-agent training set of games was significantly higher than in the no-training and all-human training set of games. Thus people who trained with two agents learned to be more cooperative. As illustrated in Figure 4.2, people's average score in the two-agent training environment (238 points) was higher than in the no-training (234 points), the single-agent training (237) and the agents-design environments (237.5). Nonetheless, this difference was not significant. We present a summary of people's behavior across all conditions in Table 4.3.

We attempt to explain the difference between the players' performance, and why certain people performed better. We distinguish between the top- and low-scoring human players in

	Follow	Stick	Heuristic	Across	Score
No Training	8	11.7	10	11.8	234.56
Training with humans	8.82	7.13	8.2	9	233.18
Training with 1 agent	10.3	11.3	10.4	10.3	237
Training with 2 agents	17	21	7.66	25.33	238.5
Training via agent design	15.7	12.64	10	15.55	237.5

Table 4.3: A summary of people's performance across all conditions in the testing session

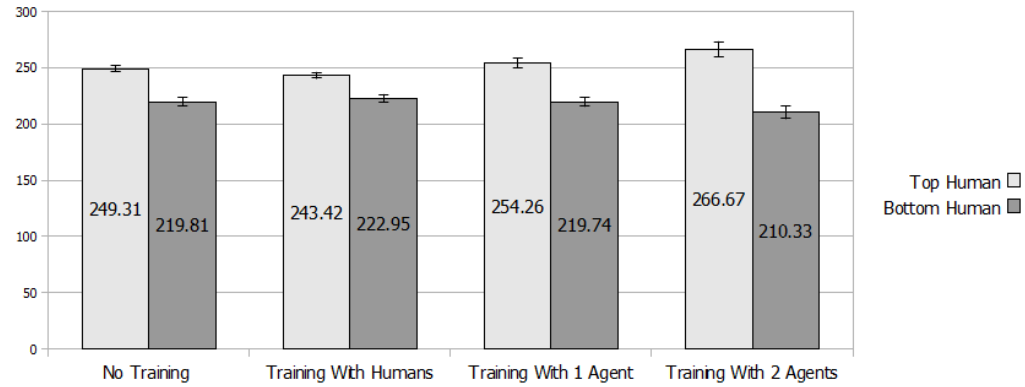


Figure 4.3: Performance comparison: People versus people

each game (the people with the highest and lowest scores each game). Because the Lemonade Stand game is constant game, one player's win is another player's loss. In the context of the Lemonade Stand game, this means that when two players coordinate and play across, they essentially earn more points than the player who is left out. We hypothesized that top-scoring human players coordinated more often with the standardized agent than low-scoring players, allowing them to outperform the low-scoring players.

Figure 4.3 shows the average performance of low- and top-scoring human players in all conditions. As depicted in the figure, the top-scoring players in the two-agent training condition outperform top-scoring players in the one-agent and all-human training conditions. In addition it is apparent that there is no difference in the performance of the low-scoring agents across the conditions. Table 4.4 shows the frequency of across outcomes in the two-agent training games for low- and top-scoring human players. As shown in the table, the top-scoring players played across moves significantly more than low-scoring players in each of the environments. In addition, top-scoring players achieved significantly more across moves in the two-agent training games than in the all-human and single-agent training games. This confirms our hypothesis, in that the success of top-scorers in the game is attributed to the players' increased coordination with the standardized agent (rather than with the other human player).

When comparing the behavior of people and of the agents, we found that in all the

	Follow	Stick	Across
low-scoring	16	15	14.92
top-scoring	**26.6**	**20**	**25.3**

Table 4.4: Number of "Follow" and "Stick" strategies used by top- and low-scoring human players in the two-agent training condition

second phase game environments the sandwich strategy was present in less than 8% of the rounds. However since sandwich is not a Nash equilibrium we did not expect two players to succeed in performing it regularly; this cooperation is not stable as the sandwiched player can move and "escape". In contrast, the "Across" method, where two players cooperate to force the third player to receive a 6 point pay-off, is a more stable form of cooperation. This form of cooperation was achieved more often by the players, specifically by the players who scored more points. Under all conditions except the two-agent training, the agent played significantly more "Stick", and "Follow" moves and thus more "Across" moves. Seemingly, the players training with two agents, learned that the "Across" move is a good strategy and were able to mimic the Training Agents' behavior, thereby increasing their score.

4.3 Cognitive hierarchy modeling of people's behavior

Cognitive hierarchy offers a model for human behavior in a competitive environment that is an alternative to Nash equilibrium [3]. It is motivated by the idea that players use strategic thinking, but are constrained by memory or time, which causes them to reach different conclusions than a Nash equilibrium [11]. According to this theory, each individual maintains a belief about the strategies of the other players, and plays his best-response based on his beliefs. In this section we will apply the K-Level hierarchy model to explain peoples' behavior in the Lemonade Stand game.

The significance of this model in multiplayer environments such as the Lemonade Stand game, where other players' actions have more influence on the outcome than the player's action, is that it is important to try to predict how other players will play, and react to this prediction in order to coordinate with the other players. The cognitive hierarchy model defines strategies in an increasing inductive order, where a level k strategy is the best response strategy to a *level k-1* strategy. Thus strategies form a hierarchy where *level 0* strategy is the lowest level strategy. Since higher level strategies require more effort to calculate it is assumed that players who play these strategies have a better understanding of the domain, thus they play better. If for example the first level (level 0) is a randomized strategy, the next step is to consider the best strategy assuming the other players use the level 0 strategy.

In a three player game such as the Lemonade Stand game two approaches to each level exist. The first approach suggests that a level K player will try to optimize its behavior while playing two level $k-1$ players. According to the second approach two level K players will work together at the expense of the third player (which is a level $k-1$ player). A model of this type was suggested by [35], where a level 0 player plays randomly with a probability x to remain in the same place two turns in a row. The explanation for using a non-uniform random strategy is that in repeated games such as this, a basic strategy could also be to stay in the same position, making the prediction of moves very simple. Once we have a level 0 strategy we can derive higher level strategies, as demonstrated in table 4.5.

The level 1 strategy would be to play across from the player who sticks (remains in the same position) with the highest probability. This is due to the fact that we can only predict the other player's move if he or she "sticks". The best response to a sticky play is to play across from his position, forcing the third player to earn 6 points while on average the two players who collaborate (play across from each other) will each earn 9 points. In a situation where 2 players play a uniform random move, The third player's move has no effect on the score and each player will earn an average score of 8.

The Level 2 strategy would be to "Stick" (stay in the same place) if we assume that there is at least one "level 1" player that will attempt to play across with the most "stickiest"

Level	Strategy
Level 0	Stick with probability x , uniform random function otherwise
Level 1	Play at the position across the player with the highest probability to "Stick"
Level 2	Stick in the same position
Level 3	Sandwich the "Stick" player with another Level 3 player

Table 4.5: Possible Cognitive hierarchy model for the Lemonade Stand game

player; a player should try to stick as much as possible to encourage such cooperation. The Level 3 strategy would be to try to "Sandwich" the level 2 players. For this level to be effective 3 players are required. Each of the level 3 players will earn 11 points and the level 2 player will only earn 2 points.

This cognitive hierarchy model of the Lemonade Stand game could explain some players' actions. When reviewing the empirical results we will attempt see how well people fit this model.

We will now explain people's behavior using the cognitive hierarchy model we introduced earlier in section 4.3. Since people can play different strategies during a single game (each strategy could be considered a different level), we assign each player a level based on his most dominant strategy. We summarize each player's moves i.e. for each move we assign the most suitable level and the player's level is deduced from his most common strategy. Each move can be assigned to either "Follow", "Stick" ,"Follow and Stick" or "Random". To determine whether a move is "Follow" or "Stick" we simply need to check the condition of the board in the previous round. The "Follow and Stick" exist to cover the scenario where two players cooperate in an "Across" strategy. For example if in round 10 of the game player A moved to position 9, and player B moved to position 3, and in round 11 both players stayed in the same location we would characterize player A's move to be "Follow and Stick". Namely, for the purpose of the analysis this move would be considered as both follow and stick. We summarize the results in Table 4.6. In this table we can see that the two groups who trained with agents and the group that designed agents suit the higher levels of this model. The results also show that training with automated agents reduces the amount of people who play at level 0. While we cannot state that people who were assigned to the level 0 group have no plan, we can assert that their patterns are harder to anticipate and therefore it is harder for other players to try to form collaborations with them. This can explain the improved performance of the groups. That is to say, playing at higher level means playing better - which leads to better performance and better results. Feedback that we received from participants who did not train or trained with people indicated that they felt the game results were random and that they could not influence their scores. The trained groups' feedback indicated more understanding of the game, where several players explained that they used the "Stick" or that they tried to form an "Across" collaboration with other players.

We will now focus on the important results that appear in Table 4.6.

- People who trained with people played level 0 significantly more than people who trained with agents.

- The group who trained with one agent had more level 2 players than the group who trained with two agents. This can be explained by the fact that playing two agents requires a player to take a more active role in finding a partner in order to succeed (play across with one agent before the two agents collaborate).

- People who trained with agents or designed agents, played level 0 strategies less than people who trained with humans or did not train. Consequently they played higher level strategies.

- After training with automated agents people played level 0 strategies significantly less, in contrast to people who trained with humans, who kept using level 0 strategies during the testing sessions.

	Level 0 - Random with probability x to stick	Level 1 - Follow	Level 2 - Stick
No training	14 (43.7%)	11 (34.3%)	7(21.8%)
Training with humans training sessions	72(72.7%)	7(7%)	20(20.2%)
Training with one agent training sessions	54(71%)	6(7.8%)	16(21%)
Training with humans test session	26(68.4%)	5(13.1%)	7(18.4%)
Training with one agent test session	15(40.5%)	6(5.2%)	16(43.2%)
Training with two agents test session	2(8.3%)	14(58.3%)	8(33.3%)
Training via agent design	9(28.1%)	14(43.7%)	9(28.1%)

Table 4.6: People's behavior according to the cognitive hierarchy model

Chapter 5

Discussion and conclusion

This thesis concerns people's behavior in a 3 player strategic environment. Initially we noticed that people had difficulty adjusting to the environment and performed poorly. To address this issue we decided to try to improve people's performance by training them by means of different methods.

In this thesis we present interesting results in the Lemonade Stand game, we try to answer the question of whether people can learn and improve their performance by using simulations with automated agents or by designing such agents. Our results show that such training is possible, and is superior to training with inexperienced people; this is true even though the domain requires players to adapt to other players in order to try to influence their behavior. It is interesting to note that people were able to learn from agents even when training with one human and one automated agent. Accordingly, people are able to learn even though they do not observe two players who have some plan of how to coordinate. The second approach of training by designing agents was also successful in improving people's performance. This method resulted in people behaving differently than the automated agent, which is expected due to not being exposed to games with other agents. We assume that people who design automated agents would analyze the game and try to formulate a plan resulting in a better chance of finding a good strategy to improve their performance. It is important to note that designing agents is a training method suited only to a small part of the population, as not everyone have programming skills. We therefore conclude that it is possible to improve people's performance by training them via automated agents, or by having them design agents. Training with agents allows people to learn new strategies by observing the other players, whereas designing agents provides them with a closer look at the game and thereby helps them gain insight about the environment in order to develop their own approach.

We also revealed an interesting heuristic people tend to use in their game. This heuristic is “move to the middle of the large gap between opponents”. This heuristic could be attributed to people's intuition from the real world, where in the real world the lemonade stand should be as far away from other lemonade stands as possible. In the Lemonade Stand game however a player will receive the same score in any position in a certain gap. If all of the players use this heuristic as their strategy, the only stable solution for the game is the “equilateral” move where each player earns 8 points.

At the end of the experiment we asked the subjects to fill in a questionnaire. The questionnaire included demographical questions, some questions about the GUI of the game and also questions about the strategy the players used. A review of the responses revealed interesting strategies such as ”If I won I stayed in my position otherwise I moved to the position of the player with the highest score”. We found that many of the players who understood the “Across” strategy tried to use the “Follow” move rather than the “Stick” move. Unexpectedly some people stated that they did not find any logic to the game and simply played randomly. Other common answers were “I tried to be unexpected”, “I tried to hurt the leading player” , “I tried to ,move to the most open spot in the circle” and lastly “I tried to sandwich stationary players”.

A number of people (19) described one of the higher levels of the Cognitive hierarchy model in their strategy. When comparing their assigned level in our experiment we found that 12 of these people were considered to play “Stick”(level 2), 5 were considered to play

"Follow"(level 1) and 2 were considered to play "Random"(level 0). Eleven players described themselves as playing completely Random, another 6 described themselves as "trying to be unexpected". We believe that these players did not understand the importance of cooperation in the game and that they either thought it was impossible to improve their results, or they thought that being expected could lead other players to take advantage of this knowledge, for example, by using the "Sandwich" strategy. Eight people described their strategy as "trying to hurt the leading player" by either standing on him or moving closer to him. We attribute this strategy to a sense of competition that the game can create among some people, who were most likely more interested in achieving the top score in comparison to the other players. The response of 'trying to go to the most open spot in the circle" is presumably a natural description of the heuristic we found people to use of going to the middle of the large gap between the opponents. This heuristic is probably a "regret minimization" method, where people try to minimize their losses by securing a certain amount of points in case at least one of the other players will not move.

5.1 Limitations

Our research contain several limitations

- Our method improved people's performance so that it was as good as the performance of standardized agent, but we were unable to train them to surpass its performance.
- We did not test how people who train by means of different methods compete against one another in an all human test group. These limitations can be solved by further research in future work.

5.2 Future work

In this work we devised a way to train people by accessible methods without the need of relying on other people to train them. In future work we plan to study the effect of training with computer agents on people playing other people. In other words we will check whether the trained people will achieve better results when playing two untrained people? We will also check whether people can outperform their own agents in a tournament, where the agents have a certain strategy which they will not deviate from and people are more flexible and can change their strategy over time. Another way of training people would be to have people watch an expert play a game while commentating and analyzing his reasoning for each move. This type of training could help people understand the logic behind each move in order to gain a better understanding of certain strategies.

Appendix A

The Lemonade Stand game

A.1 instructions and quiz for the Lemonade Stand game

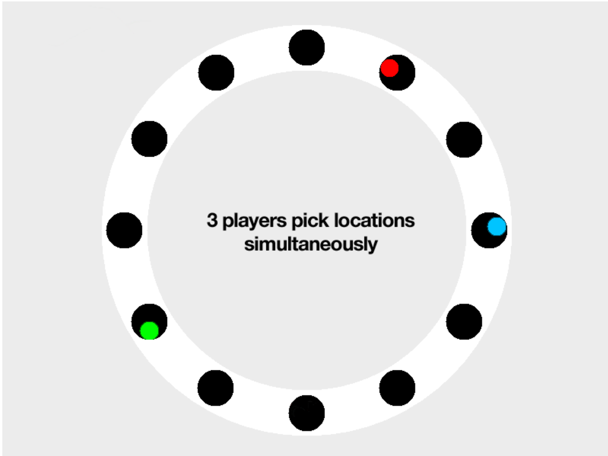
Lemonade Stand Game Instructions

Note: If you have already performed a HIT involving this game, please do not play again. You will not be paid for the second time.

Note: You must have java installed in order to participate in this task, if you can't see the picture bellow than you need to download and install java.

3 Players, 12 Locations

In this game, you and two other players must set up lemonade stands on a circular island. There are twelve locations along the beach where you may set up your stand. To choose where to place your lemonade stand, click on one of the locations.

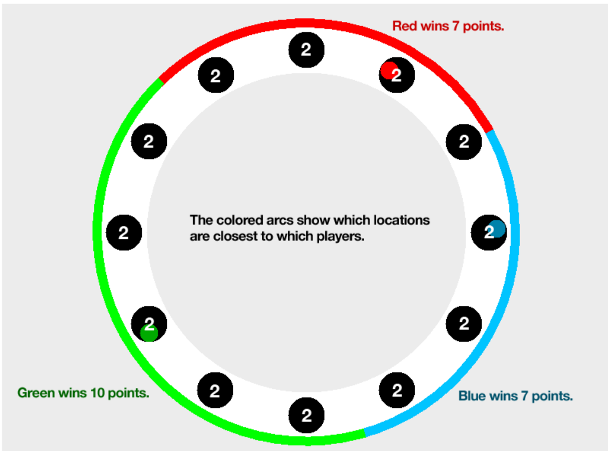


Scoring

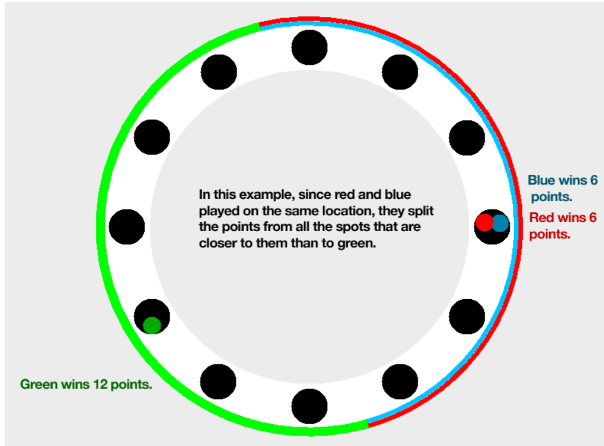
There are two customers at each spot. Once everyone has placed their lemonade stand, all the customers will go to the nearest stand and buy lemonade. For every customer who visits you, you get one point.

An equivalent way to calculate your score is to add the number of locations between you and the player to your right (including the location that player is on) and the number of locations between you and the player to your left.

If two stands are equally far away from a spot, they will split the customers from that spot.



If two players play in the same location, they split the points from all the locations nearest to them.

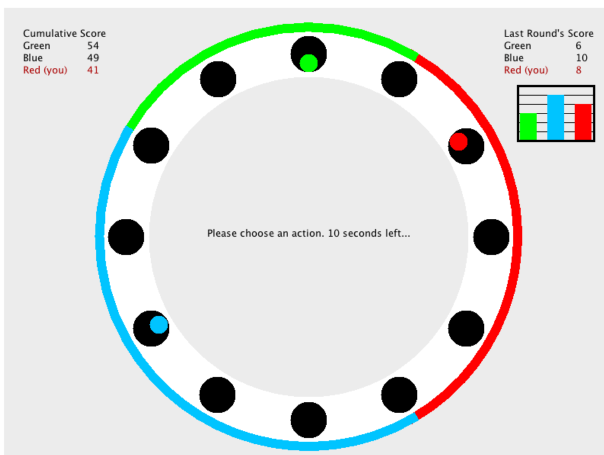


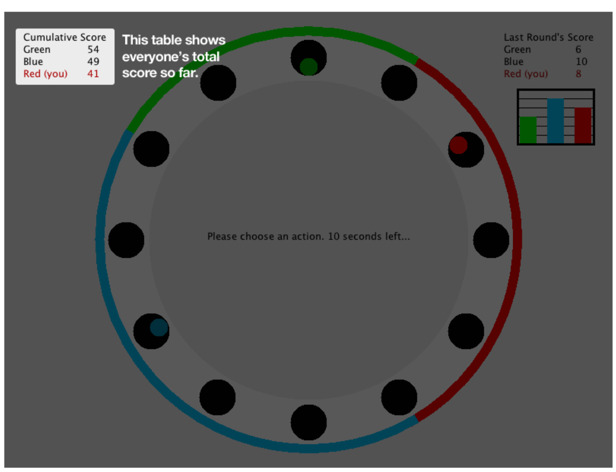
If everyone plays in the same location, everyone gets 8 points, since there are 24 points on the whole board on every turn.

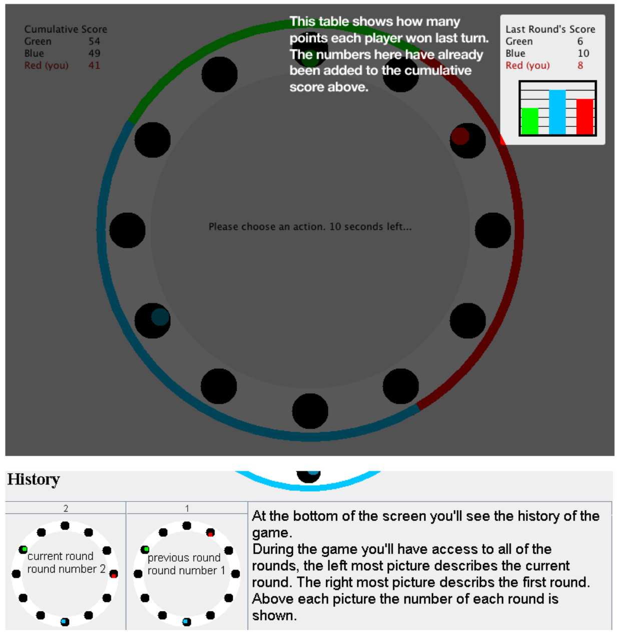
Try to maximize your score

The game will last for 30 turns. You should try to get as high a score as you can.

Here is an example of what your screen will look like while you play the game. In this example, you are the red player. In the actual game, you might end up playing as blue or green instead.



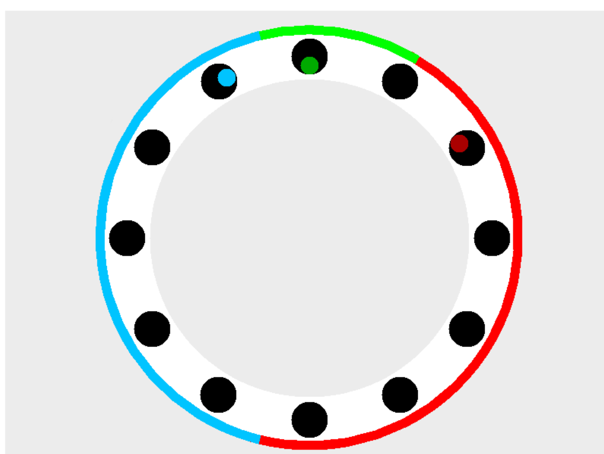




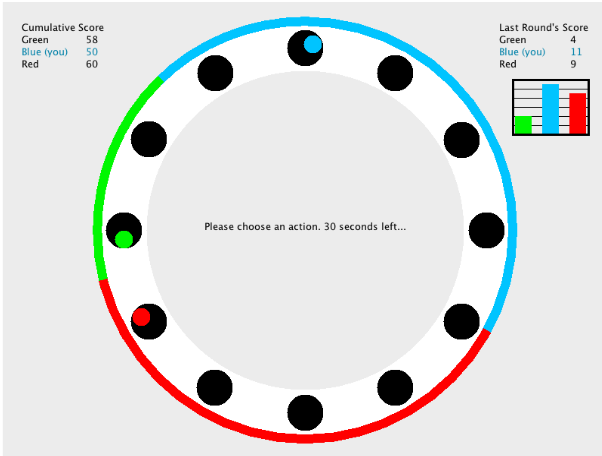
Quiz

Please answer both questions correctly to proceed.

You must answer the following questions correctly in order to play the game.



1. In the above picture, how many points did the green player score?
   - 0
   - 1
   - 1.5
   - 3
   - 6

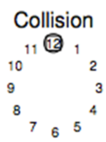


2. In the above picture, how many points have you scored in total so far this game?
- ○ 4
- ○ 9
- ○ 11
- ○ 50
- ○ 58
- ○ 60

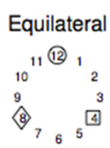
A.2 Advance strategies presented to the subjects

Lemonade Stand Game Basic Strategies

In this game, you and two other players must set up lemonade stands on a circular island. There are twelve locations along the beach where you may set up your stand. To choose where to place your lemonade stand, click on one of the locations. Stay in one place for consecutive turns, staying in the same place will encourage other players to stay away from that place, and even be considered as an invitation to the Across strategy(see later). However this could lead to exploitation in the form of a sandwich move (see later).



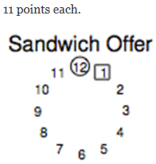
Two players on the same spot form a collision, this is primarily bad, and both players will only earn 6 points. However this could be used to push other players from certain location.



All three players earn 8 points.

If two players play the across (opposite sides), they force the third player to earn exactly 6 points, and split the remaining 18 between them.

Two players sandwiching a player, the middle player will only earn 2 points, while the sandwiching players earn 11 points each.



An offer for a sandwich, this could be used on a player if you can anticipate his next move, this could signal the third player that you are interested in trying to form a sandwich.

Bibliography

[1] H. Aguinis and K. Kraiger. Benefits of training and development for individuals and teams, organizations, and society. *Annual review of psychology*, 60:451–474, 2009.

[2] A. Byde, M. Yearworth, K.Y. Chen, and C. Bartolini. Autona: A system for automated multiple 1-1 negotiation. In *E-Commerce, 2003. CEC 2003. IEEE International Conference on*, pages 59–67. IEEE, 2003.

[3] C.F. Camerer, T.H. Ho, and J.K. Chong. A cognitive hierarchy model of games*. *Quarterly Journal of Economics*, 119(3):861–898, 2004.

[4] M. Campbell, A.J. Hoane, and F. Hsu. Deep blue. *Artificial intelligence*, 134(1):57–83, 2002.

[5] J.A. Colquitt, J.A. LePine, and R.A. Noe. Toward an integrative theory of training motivation: a meta-analytic path analysis of 20 years of research. *Journal of applied psychology*, 85(5):678, 2000.

[6] JJ Cromby, P.J. Standen, D.J. Brown, et al. The potentials of virtual environments in the education and training of people with learning disabilities. *Journal of Intellectual Disability Research*, 40(6):489–501, 1996.

[7] D. Druckman and N. Ebner. Onstage or behind the scenes? relative learning benefits of simulation role-play and design. *Simulation & Gaming*, 39(4):465–497, 2008.

[8] S.L. Epstein. Toward an ideal trainer. *Machine Learning*, 15(3):251–277, 1994.

[9] I. Erev and A.E. Roth. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American economic review*, pages 848–881, 1998.

[10] Y. Gal, A. Pfeffer, F. Marzo, and B.J. Grosz. Learning social preferences in games. In *Proceedings of the National Conference on Artificial Intelligence*, pages 226–231. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2004.

[11] Ya'akov Gal, Barbara Grosz, Sarit Kraus, Avi Pfeffer, and Stuart Shieber. Agent decision-making in open mixed networks. *Artificial Intelligence*, 174(18):1460–1480, 2010.

[12] Yaakov Gal and Avi Pfeffer. A language for opponent modeling in repeated games. In *AAMAS Workshop on Decision-Theoretic and Game-Theoretic Agents*. Citeseer, 2003.

[13] Ece Kamar, Eric Horvitz, and Chris Meek. Mobile opportunistic commerce: mechanisms, architecture, and application. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 2*, pages 1087–1094. International Foundation for Autonomous Agents and Multiagent Systems, 2008.

[14] Patrick Kenny, Arno Hartholt, Jonathan Gratch, William Swartout, David Traum, Stacy Marsella, and Diane Piepol. Building interactive virtual humans for training environments. In *The Interservice/Industry Training, Simulation & Education Conference (I/ITSEC)*, volume 2007. NTSA, 2007.

[15] H.J. Klein, R.A. Noe, and C. Wang. Motivation to learn and course outcomes: The impact of delivery mode, learning goal orientation, and perceived barriers and enablers. *Personnel Psychology*, 59(3):665–702, 2006.

[16] S. Kraus and D. Lehmann. Designing and building a negotiating automated agent. *Computational Intelligence*, 11(1):132–171, 2007.

[17] R. Lin, Y. Gal, S Kraus, and Y. Mazliah. Training with automated agents improves peoples behavior in negotiation and coordination tasks. *Decision Support Systems*, 2012.

[18] R. Lin, Y. Gev, and S. Kraus. Bridging the gap: Face-to-face negotiations with an automated mediator. *IEEE Intelligent Systems*, 26(6):40–47, 2011.

[19] R. Lin and S. Kraus. Can automated agents proficiently negotiate with humans? *Communications of the ACM*, 53(1):78–88, 2010.

[20] Y. Mazliah and Y. Gal. Coordination in multi-player human-computer groups. In *First International Workshop on Human-Agent Interaction Design and Models*, 2012.

[21] R. R. Murphy. Humanrobot interaction in rescue robotics. *IEEE Transactions On Systems, Man, and CyberneticsPart C: Applications and Reviews*, 34(2), May 2004.

[22] C. Nass and B. Reeves. The media equation: How people treat computers, televisions, and new media as real people and places, 1996.

[23] D. Rajarshi, J. E. Hanson, J. O. Kephart, and G. Tesauro. Agent-human interactions in the continuous double auction. In *Proc. 17th International Joint Conference on Artificial Intelligence (IJCAI)*, 2001.

[24] A. Rimmel, O. Teytaud, C.S. Lee, S.J. Yen, M.H. Wang, and S.R. Tsai. Current frontiers in computer go. *Computational Intelligence and AI in Games, IEEE Transactions on*, 2(4):229–238, 2010.

[25] FD Rose, EA Attree, BM Brooks, DM Parslow, and PR Penn. Training in virtual environments: transfer to real world tasks and equivalence to real task training. *Ergonomics*, 43(4):494–511, 2000.

[26] N. Schurr, P. Patil, F. Pighin, and M. Tambe. Using multiagent teams to improve the training of incident commanders. In *Proc. 5th International Joint Conference on Autonomous Agents and Multi-agent Systems (AAMAS)*, 2006.

[27] B.M. Slator, P. Juell, P.E. McClean, B. Saini-Eidukat, D.P. Schwert, A.R. White, and C. Hill. Virtual environments for education. *Journal of Network and Computer Applications*, 22(3):161–174, 1999.

[28] A.M. Sykulski, A.C. Chapman, E.M. de Cote, and N.R. Jennings. Eaˆ 2: The winning strategy for the inaugural lemonade stand game tournament. In *ECAI 2010: 19th European Conference on Artificial Intelligence*, pages 209–214, 2010.

[29] D.L. Tate, L. Sibert, and T. King. Using virtual environments to train firefighters. *Computer Graphics and Applications, IEEE*, 17(6):23–29, 1997.

[30] D. Traum, S. Marsella, J. Gratch, J. Lee, and A. Hartholt. Multi-party, multi-issue, multi-strategy negotiation for multi-modal virtual agents. In *Intelligent Virtual Agents*, pages 117–130. Springer, 2008.

[31] A. van Wissen, Y. Gal, B. Kamphorst, and V. Dignum. Human-agent team formation in dynamic environments. *Computers in Human Behavior*, [to appear].

[32] D. Waller, E. Hunt, and D. Knapp. The transfer of spatial knowledge in virtual environment training. *Presence*, 7(2):129–143, 1998.

[33] P.L.T. Weiss, Y. Naveh, and N. Katz. Design and testing of a virtual environment to train stroke patients with unilateral spatial neglect to cross a street safely. *Occupational Therapy International*, 10(1):39–55, 2006.

[34] M. Wunder, M. Kaisers, M. Littman, and J.R. Yaros. A cognitive hierarchy model applied to the lemonade game. In *AAAI Workshop on Interactive Decision Theory and Game Theory (IDTGT)*, 2010.

[35] M. Wunder, M. Kaisers, M. Littman, and J.R. Yaros. A cognitive hierarchy model applied to the lemonade game. In *AAAI Workshop on Interactive Decision Theory and Game Theory (IDTGT)*, 2010.

[36] M. Zinkevich. The lemonade game competition `http://tech.groups.yahoo.com/group/lemonadegame/`, 2009.

[37] M.A. Zinkevich, M. Bowling, and M. Wunder. The lemonade stand game competition: solving unsolvable games. *ACM SIGecom Exchanges*, 10(1):35–38, 2011.