

Modeling Agents Based on Aspiration Adaptation Theory

Avi Rosenfeld · Sarit Kraus

Received: date / Accepted: date

Abstract Creating agents that realistically simulate and interact with people is an important problem. In this paper we present strong empirical evidence that such agents should be based on bounded rationality, and specifically on key elements from Aspiration Adaptation Theory (AAT). First, we analyzed the strategies people described they would use to solve two relatively basic optimization problems involving one and two parameters. Second, we studied the agents a different group of people wrote to solve these same problems. We then studied two realistic negotiation problems involving five and six parameters. Again, first we studied the negotiation strategies people used when interacting with other people. Then we studied two state of the art automated negotiation agents and negotiation sessions between these agents and people. We found that in both the optimizing and negotiation problems the overwhelming majority of automated agents and people used key elements from AAT, even when optimal solutions, machine learning techniques for solving multiple parameters, or bounded techniques other than AAT could have been implemented. We discuss the implications of our findings including suggestions for designing more effective agents for game and simulation environments.

We would like to acknowledge the many productive conversations and visits with Prof. Reinhard Selten, Dr. Martin Hohnisch, and Sabine Pittnauer of the University of Bonn. We would also like to thank the anonymous reviewers for their comments, as well as Yael Ejgenberg, Ron Adany, Yosi Ben-Agu, Roni Toledano, Anat Sevet, and Yael Blumberg for judging the agents in Section 4. This work was supported in part by the National Science Foundation under grant #0705587. Sarit Kraus is also affiliated with UMIACS. A preliminary version of this paper with a subset of these results was published in IJCAI 2009.

Avi Rosenfeld
Department of Industrial Engineering of the Jerusalem College of Technology
Jerusalem, Israel 91160
E-mail: rosenfa@jct.ac.il

Sarit Kraus
Department of Computer Science of Bar-Ilan University
Ramat-Gan, Israel 52900
E-mail: sarit@cs.biu.ac.il

1 Introduction

The challenge of realistically modeling individual reasoning and decision making is of utmost importance for many economics and artificial intelligence researchers [1, 2, 8, 13–16, 22–25]. Within the field of economics, validly encapsulating human decision-making is critical for predicting the short and long term effects of a given policy. For computer scientists, this is critical for mixed human-computer systems such as entertainment domains [13], Interactive Tutoring Systems [16], and mixed human-agent trading environments [14]. In these and similar domains, creating agents that effectively understand and/or simulate people’s logic is particularly important. To date, these agents have been traditionally created based on the perspectives of unbounded rationality including expected utility, game theory, Bayesian models, or Markov Decision Processes (MDP) [17, 21].

While these models are mathematically elegant and have proven effective in some applications [4, 17, 21, 30], several fundamental obstacles exist in applying them to many real-world problems. First, previous research in experimental economics and cognitive psychology has shown that human decision makers often do not adhere to fully rational behavior. For example, Kahneman and Tversky [5] have shown that individuals often deviate from optimal behavior as prescribed by Expected Utility Theory. Second, decision makers often lack complete information and thus do not necessarily know the quantitative structure of the environment in which they act. Consequently, even assuming people act rationally, they cannot always compute the optimal solution for a given problem, as they lack facts required to arrive at this decision. Finally, even if people wish to act rationally, and have complete information about a given problem, it may still be impossible for them to compute the optimal solution. Previous research has found that many classes of real-world problems exist for which finding the optimal sequence of actions is of intractable computational complexity [20]. Thus, even in the best of circumstances, expecting people to behave optimally based on full rationality is unrealistic.

However, we also do not claim that people typically act irrationally, and thus a third possibility must be developed between the poles of classic rationality and irrationality. We believe that models based on Bounded Rationality hold the most promise to accomplish this achievement. This research direction, initiated by Simon [27], assumes that people – except in the most simple situations – lack the cognitive and computational capabilities to find optimal solutions. Instead they proceed by searching for non-optimal alternatives to fulfill their goals. Simon coined the term “satisfice” to capture that bounded decision makers seek “good enough” solutions and not optimal ones. In this tradition, Sauermann and Selten proposed a framework called Aspiration Adaptation Theory (AAT) [22, 23] as a model of boundedly rational decision making.

This paper presents an extensive empirical study of three different types of experiments in optimization and negotiation domains. We first studied two relatively simple optimization problems where optimal solutions required solving problems with one and two parameters. In both of these problems we asked a group of nearly 70 people to describe what strategies they would use to solve these optimization problems. We also asked a second group of nearly 100 people to write computer agents to solve these problems on their behalf. These two types of experiments focused on how people explicitly quantify their own strategies, either in plain text without any need to implement them, or in computer agents designed to solve these problems on their behalf. A third type of experiment studied people’s and agent’s behavior in two negotiation realistic scenarios with five and six parameters. In this domain we first studied human negotiation strategies by studying over 100 negotiation sessions between people. We also studied the behavior of two state of the art automated negotiation

agents [11, 12, 19] in the same two negotiation problems as they attempted to reach effective agreements with over 100 people. In these studies we reviewed the negotiation performance logs to implicitly derive key characteristics of the negotiation strategies used. This study then compared the negotiation strategies used by the expert agents versus those of people negotiating with each other. A full description of the research methodology used, as well as specifics of these problems, can be found in Section 3 of this paper.

In both domains and in both types of problems we found strong empirical evidence that these agents generally act consistently with bounded rationality, and specifically in accordance with key elements of Aspiration Adaptation Theory (AAT) [22, 23]. Section 4 of this paper details these results, and demonstrates that optimal methods and other bounded rationality theories typically were not used by the people and agents we studied. This realization allowed us to effectively model the decisions by these agents. Additionally, in Section 5 we present several guidelines by which more realistic or effective agents can be built based on this result. We begin by presenting the basics of AAT and how this theory can be generalized in the next section.

2 Aspiration Adaptation Theory

Many real-world problems cannot be easily quantified based on traditional expected utility models [17]. In our lives we must make many difficult decisions. Consider the following specific example. Assume you need to relocate and choose a new house to live in. There are many factors that you need to consider, such as the price of each possible house, the distance from your work, the neighborhood and neighbors, and the schools in the area. How do you decide which house to buy? Theoretically, utility based models can be used. However, many of us do not create rigid formulas involving numerical values to weigh trade-offs between each of the search parameters. Given this, what alternative can effectively model our decisions? In a more radical example, assume someone offered you 10,000 Euro in exchange for the last month of your life. Would you accept? In this case, utility theory seems even less appropriate due to incomplete knowledge. While you might be able to potentially evaluate the impact of 10,000 Euro today, you cannot possibly know what quality of life you can expect in your last month.

Our goal is to develop a realistic model for human decision making such that two types of computer agents can be created. As described above, many real-world problems do not lend themselves to utility theory. By better understanding how people behave, we can create agents that better interact with people in these types of problems [13, 14, 16]. For example, assume you wish to create an agent to help a user choose which house to bid on. By understanding how people generally make decisions, we believe that better agents can be created to interact with a given user. Second, by better understanding people's reasoning, we believe that we can better simulate their behavior in a variety of training or simulation environments. In these types of environments, the goal is not to interact with just one given user, but instead to accurately simulate a range of people's behavior. Therefore, by better understanding how people make their decisions, we can better simulate their behavior. In creating both types of agents, we believe that bounded rationality models, and specifically using key elements from AAT, should be used to form the basis of agents' decision making models.

Aspiration Adaptation Theory was proposed by Selten as a general economic model of boundedly rational behavior [22, 23]. As our goal is to study the applicability of this theory to computer agents, we frame this theory within the specifics of the optimization and

negotiation problems presented in Section 3. We refer the reader to their full paper [23] for a complete and general presentation of this theory from an economic perspective.

AAT is an approach by which bounded agents address a complex problem, \mathcal{G} . The complexity of \mathcal{G} prevents the problem from being directly solved, and instead an agent creates m goal variables G_1, \dots, G_m as a means for solving \mathcal{G} . These goals are assumed to be incomparable, and no aggregate utility function, known as a substitution rate in economics, can be constructed for combining the m goals into \mathcal{G} . The lack of a utility function may be due to the fact that the problem is simply too complex to quantify such a function, or because the agent lacks the resources to properly calculate the aggregate utility function for \mathcal{G} . In attempting to reach a solution, the agent has a group of s instrument variables which represent base actions that can be used to pursue the goal variables. We define a *strategy* $x = (x_1, \dots, x_s)$ as a combination of instrument values for the goal variables G_1, \dots, G_m . These values can and do typically change over the life of the agent. For example, assume a company has a goal \mathcal{G} to be optimally profitable. Goal variable examples might be to create more sales, create higher brand awareness or to invest in the company's production infrastructure. Here, the instrument variables may include lowering the product's price, investing more money in marketing, or hiring better skilled workers. The agent might have one strategy at the beginning of its operation (e.g. an opening sale to entice buyers) and then use a different strategy after the business matures. An action A is a rule for changing the strategy. A finite number of n actions, A_1, \dots, A_n are considered. The agent can only choose one action per time frame.

Examples of actions in this context include:

- Raising the product's price by 5%.
- Lowering the product's quality by 10%
- Lowering the product's price by 10% in conjunction with raising the marketing expenditure by 15%
- Making no change to the strategy

Despite its lack of utility to quantify rational behavior, the model provides several guidelines for how bounded agents will behave. Agents decide about goal variables as follows: The m goal variables are sorted in order of priority, or the variables' *urgency*. Accordingly, the order of G_1, \dots, G_m refers to the goals' urgency, or the priority by which a solution for the goal variables should be attempted. Each of the goal variables has a desired value, or its *aspiration level*, that the agent sets for the current period. Note that this desired value is not necessarily the optimal one, and the agent may consider the variable "solved" even if it finds a sub-optimal, but yet sufficiently desired value. The agent's search begins with an initial aspiration level and is governed by its *local procedural preferences*. The local procedural preferences prescribe which aspiration level is most urgently adapted upward if possible, second most urgently adapted upward if possible, etc., and which partial aspiration level is *retreated from* or adapted downward if the current aspiration level is not feasible. Here, all variables except for the goal variable being addressed are assigned values based on *ceteris paribus*, or all other goals being equal, a better value is preferred to a worse one.

For example, assume that a company needs to be optimally profitable by setting goal variables relating to a product's price and quality. While the company's head might want to both charge a high price and provide high quality, market conditions will likely prevent this. This will typically occur as people will not buy the product at too high a cost and the investment in quality might too strongly impact the company's profit margin. Thus, the company's decision maker might form an *urgency* between the price and quality variables and an *aspiration level* for what constitutes a "good enough" value for these variables. Borrowing from Simon's terminology [27] there is only an attempt to "satisfice" the goal values, or achieve

“good enough” values for these variables instead of trying to optimize or truly solve them. The variables’ *urgency* dictates which goal is attempted to be satisfied first. In this example, let us assume the product’s price has the highest urgency. As such, only once the decision maker is satisfied by the value for this goal will she proceed to the second variable, in this case, quality. However, if it is not possible to achieve all aspirations, the original values will need to be reexamined, or retreated from. Thus, the agent may retreat, or rethink raising the product’s price and raise its quality instead. Similarly, referring to the above example of searching for a new home, one would like find the biggest home, all other goals being equal. However, if this is found to be impossible, the agent may retreat from, or settle for a smaller home, but satisfy another goal, such as the home’s location.

The search procedure as described by AAT is different from traditional search methods, such as Hill-climbing or machine learning methods such as Gradient Descent techniques, in two aspects. First, with traditional learning or search, optimal values for all variables are sought simultaneously [21]. In contrast, within AAT only one goal is attempted to be satisfied for at a time. Second, in AAT, the focus is on “satisficing” goal values based on their *aspiration levels*. This approach makes no attempt to find optimal values beyond these “good enough” values – something machine learning methods do search for.

While this theory has existed since the early 1960’s [22], there are few empirical studies of how well it explains observed behavior [26]. As Selten’s paper states [23], “AAT was meant as a useful point of departure for theory construction. The theory as it stands cannot claim to be a definite answer to the problem of modelling bounded rationality. Probably, one would need extensive experimental research in order to modify it in a way which fits observed behavior.” In this paper, we present strong empirical evidence that several key elements of AAT effectively model computer agent’s and people’s behavior in dealing with two types of problems within the optimizing and negotiation domains. To the best of our knowledge, this paper represents the first study that bridges between the fields of experimental economics and computer science to demonstrate the applicability of bounded rationality theory in modeling these agents.

It is important to note one key difference between AAT as originally formulated, and the decision making process in the problems we studied. AAT was formulated when no utility could possibly be constructed, such as the example presented above of trading 10,000 Euro for one month of one’s life. However, in the problems we considered some utility could be constructed, albeit with some difficulty, and more closely resembles cases such as house purchases or setting the price or quality of a product. As such, one may assume that AAT is not relevant, and other models, bounded or utility based, should be used instead.

In fact, we found that many key elements of AAT can still be applied in many cases where utility can be used; however, people shy away from such models due to the complexity of the problem they must solve. Thus, this paper represents a significant generalization of AAT theory and explores some consequences of this finding in creating better and more realistic agents. To demonstrate this point, we studied a total of four problems, two relatively simple optimization problems, and two real-world negotiation problems. In all cases we compared the decision models of people and agents operating within these problems to AAT, other bounded rationality models, and optimal utility based methods. The next section details the exact methodology used in this study.

3 Research Methodology, Motivation, and Domain Descriptions

The basis of our methodology is to study agents designed to operate within a series of progressively more complex problems. Our empirical study begins with studying relatively simple optimizing problems involving one and two parameters, and concludes with studying realistic negotiation problems involving five and six parameters. By studying less complex problems first, we believe it will be easier to understand the general decision making models used by people and those used by agents programmed to operate on their behalf. By focusing on more complex problems as well, we show the applicability of our findings to real-world applications allowing us to suggest future directions.

We specifically studied negotiation problems as various real-world tasks are based on negotiation capabilities. These can be as simple and ordinary as haggling over a price in the market or deciding what television show to watch. Negotiation issues can also involve issues where millions of lives are at stake, such as interstate disputes [32] and nuclear disarmament [3]. No matter what the domain, the negotiation process itself is not an easy task. The parties will have conflicting interests with reference to some aspects of the negotiation. On the other hand, both sides might also have the incentive to cooperate with each other, as reaching an agreement could be more beneficial for them than walking away without any agreement [18]. The use of simulation and role-playing is common for training people in negotiations (e.g., the Interactive Computer-Assisted Negotiation Support system (ICANS) [29]¹, the InterNeg Support Program for Intercultural REsearch (INSPIRE) [7] and virtual humans for training [6]). These simulations can be used in conjunction with people to alleviate some of the efforts required of people during negotiations and also assist people that are less qualified in the negotiation process. Additionally, there may be situations in which simulations can even replace human training procedures. Furthermore, while we studied negotiation scenarios, the results, and the lessons that we draw from them, are likely to be applicable to other mixed-human domains such as entertainment domains [13], Interactive Tutoring Systems [16], and mixed human-agent trading environments [14].

In both the optimization and negotiation domains we performed two types of experiments: in the first type of experiments we studied how people solved the problems without any need to program their solution, and in the second type of experiments we studied the computer agents people encoded to solve the same problems. The goal of the first set of experiments was to positively identify the general decision model people use. This step involves experiments exclusively with people, and allows them to choose any strategy without implementation concerns. In the optimization problems, people directly described their strategies in plain text. In the negotiation experiments, people negotiated with each other via a web interface and we studied the strategies based on performance logs of their negotiation sessions. In both domains, we then proceeded to ascertain if these strategies are consistent with key elements of AAT or other possibilities. In the second set of experiments within the two optimization problems, we studied the agents that moderately skilled programmers wrote for two relatively simple optimization problems. In the more complex negotiation problems, we studied expert automated negotiation agents. The goal was to observe which elements, if any, of AAT are present in these agents. Specific to the negotiation domain, we also discuss the lessons that can be drawn in order to create more effective expert agents. Note that previous research from economics and computer science [1,2,8,13–16,22–25] also shares the goal of learning from human behavior in order to create more effective agents.

¹ More about this exciting application can be found at: <http://www.icons.umd.edu/>.

The two sets of optimization experiments focus on the decision models of people and their agents. We recognize that quantifying these models, even those operating within relatively simple optimization problems, can be difficult. To facilitate this task, we constructed a short list of several questions to analyze these strategies. These questions included: How many instrument variables were solved for? What were they? Was a search process used to set these instrument variables, or was a predefined strategy, independent of actual performance used instead? If search was used, were the variables searched for simultaneously, or sequentially? If sequential search was used, did the agent revisit variables after it had originally set a value for it (such as to retreat, or revise downward the previously set threshold value). In the first type of optimization experiments, no programming was required, and these questions were answered based on the plain text or pseudocode people wrote to explain their strategies. In the second set of experiments, these questions could be directly answered by the agents' programming code or based on the agents performance as gleaned from log files of the agents' performance. In both cases it should be noted that these questions are general, and can be applied to any domain, regardless of the particularity of the problem characteristics. However, despite the wealth of written information and log files provided by the people and their agents, at times some ambiguity may exist in an agent's model. To overcome this issue, we had a total of five people judge each of the agents in these results. Of these five people, four were not authors of this paper, and thus had no bias. Furthermore, we studied the significance in the deviations in the answers provided by these 5 judges, to confirm that no significant differences exist. This allowed us to ascertain that these judges' decisions were truly representative of the agents' strategies.

In the second, more complex negotiation domain, we focused on the agents' behavior instead of the decision model used. The actions of the agent can be quantified easier than their decision models. For example, in checking if simultaneous negotiation was used in the six parameter negotiation problem, one could simply check the activity logs of that agent and observe how many parameters were changed in a given negotiation interaction. Similarly, if a predefined strategy was used to set a given parameter, it should become obvious over the course of hundreds of negotiation interactions that the agent does not change its behavior, regardless of the specifics of a given problem. Thus, this analysis may be more accurate than relying exclusively on judges to determine an agent's decision model.

When categorizing the agent's decision model or actions, both AAT and many other possibilities may theoretically be a proper fit. In both problems in the two domains, optimal search methods or negotiation strategies could have been used. Even if optimal strategies are not used, several different types of bounded strategies other than AAT are possible. Psychological models of bounded rationality have suggested that people use domain specific biases or heuristics to solve problems sub-optimally [2,5]. These methods are different from bounded economic approaches such as AAT which present a general, non-domain specific method for searching for a satisficing solution. Following the psychological approaches, simple predefined or greedy heuristics could be used to set values for the unknown parameters. In the optimizing domains, predefined values could have been used instead of attempting to solve the one and two parameter problems. In the negotiation domains simple compromise heuristics could have been used. Possibilities of such heuristics include always countering the middle position between the previous offer of both sides or offering the middle position between all previous offers of both sides. These types of approaches would be consistent with basic implementations of Gigerenzer and Goldstein's fast and frugal heuristics [2] and involve using simplistic preset values that are seen as "good enough". Other non-AAT search approaches may involve simultaneously searching for all unknown values. More traditional A.I. methods such as Hill-climbing or Gradient Descent [21] can

simultaneous search for multiple variables, something not consistent for AAT's concept of urgency.

3.1 Optimization Problems

In this paper we first study the strategies people, and the agents they write on their behalf, use within two optimization problems – a commodity search problem where the optimal solution could be found based on solving one cost instrument variable, and a more complicated domain where the optimal solution requires solving for price and quality instrument variables. In both problems an optimal solution can be constructed, and thus bounded rationality theories such as AAT are potentially unnecessary. However, the optimal solution is far from trivial in these domains. Consequently, these problems allow us to explore issues surrounding the strategies and heuristics used.

3.1.1 Research Methodology for Optimization Problems

The strategy-method from experimental economics [8,24,25] is central to the methodology of studying the optimization problems. The strategy method is typically implemented in two stages. First, people are encouraged to fully study a problem at hand. Next, after they have a complete understanding of the task, they are then able to formulate a complete strategy for that problem. The assumption behind this method is that people can effectively describe their own strategies within a certain period of time, without additional aids such as handbooks or other sources of information. Underlying this assumption is that people will execute a task to the best of their abilities if they are properly motivated, monetarily or otherwise.

In our study, we asked people to formulate their strategies in one of two ways. In the first variation we took a group of 50 people ranging in ages from 18 to 65 with a nearly 50% distribution between men and women. We presented them with written instructions about the optimization problem to be solved. We then allowed them to freely use a web-based simulation to fully understand the optimization problem at hand. Finally, we asked them to write what strategy they used to solve the problem. In this variation, the strategy method involves formulating strategies as text or pseudo-text, and does not require the participants to use programming languages to encapsulate their ideas. This study follows previous experimental economics experiments [8,24] where participants did not require any programming knowledge. In contrast, in the second study, we further studied not only the decision making process of people, but the performance of the agents they wrote to search for an optimal solution on their behalf. As people familiar with programming concepts were needed, we limited participants of this study to upper level undergraduates (seniors), masters and PhD computer science students ranging in age from 18 to 30 with a nearly uniform distribution between men and women, and asked them to write their agents within a firm 2 week deadline. The undergraduate students had a minimum of two years programming experience, and the PhD students had a minimum of 5 years of programming experience. As motivation, we told the subjects that their grade was based on their agent's performance. Specifically, the performance of each agent constituted approximately 20% of the students' grade in an upper level elective on writing agents. The students' grades were determined based on a linear relationship of the agents' performance with the performance closest to the class average receiving a 75, and the best performing agent(s) being given a 100. Thus, students could guarantee a grade of 100 if they could find the optimal or even relatively best solution.

Once these programs were written, the decision model of the person’s agent was directly evaluated from the program’s logic. The approach whereby participants are expected to program their own strategies has also been used in experimental economics [25], where participants with programming knowledge were also asked to program their strategies in Pascal. This approach has also begun to be used in artificial intelligence research as well [1].

We took several steps to explore the validity behind the strategy methods assumption in the programming experiments. First, we compared the written strategies from the first set of experiments to those of agents programmed by the second group of participants. If strong similarities exist between these two groups of strategies we can confidently claim that requiring people to program their strategy does not affect the results. Additionally, in order to ensure that people were able to effectively program their own strategies, several steps were taken. First, we made sure the subjects were given a well designed Java framework which included methods for most computations (e.g. finding the average of the agent’s past performance). Thus, full knowledge of Java was not necessary, as the provided framework enabled people to encode their strategies in only a few lines of code. This approach mimics the strategy-method variant previously used in economics [25] where people were provided a programming framework to implement their own strategies. The Java language was chosen as all subjects had experience using this programming language in previous multiple courses and thus were fluent in this language. Finally, after the first two week deadline, we required that each subject submit a “draft” agent. We reported back to all of the subjects with the relative performance of their agent to others in the group. The subjects were then allowed an additional week to fix bugs in their agents, or to improve their implementation without any penalty.

We used two tools to study the agents’ design. First, we ran the agents and studied the agents’ decisions. By analyzing the agents’ logic and comments, one can often understand the search process used. Additionally, after the participants had completed programming and debugging their agents, we distributed questionnaires to them with directed questions about the strategy their agent used, confirming particulars of their approach. This allowed us to confirm what strategies were used by each of the agents.

3.1.2 One-Parameter Optimization Domain

In the first optimization problem, we consider a problem where a person must minimize the price in buying a commodity (a television) given the following constraints. In this problem, a person must personally visit stores in order to observe the posted price of the commodity. However, some cost exists from visiting additional stores. We assume this cost is due to factors such as an opportunity cost with continuing the search instead of working at a job with a known hourly wage. For any given discrete time period, the person must decide if she wishes to terminate the search. At this point, we assume she can buy the commodity from any of the visited stores without incurring an additional cost. The goal of the agent is to minimize the overall cost of the process which is the sum of the product cost and the aggregated search cost.

From a strategic point of view, the game is played under a time constraint. An optimal solution to this problem can be found as an instance of the Pandora’s problem [31] resulting in a stationary threshold below which the search should be terminated. Formally, we can describe this problem as follows:

We assume that there is a finite timeline $\mathcal{T} = \{1, 2, \dots, k\}$. In each time step $t, t \leq k$ the agent observes a cost and needs to decide whether to end the search. All of the observed

costs, regardless of the time step, are drawn from the same distribution. We denote c_t as the lowest price the agent observed up to and including the time period t (i.e., $c_t \leq c_{t-1}$). At the end of the game the agent's cost is $cost(t, c_t) = c_t + \lambda * t$, $\lambda > 0$. The agent's goal is to minimize this cost. As has been previously proven, the optimal strategy in such domains is as follows: \bar{c} exists such that if $c_t \leq \bar{c}$ the agent should stop the search [31].

Intuitively, it seems strange that the decision as to whether the agent should stop the search does not depend on how much time is left, i.e., \bar{c} does not depend on $k - t$. However, the reason for this is as follows. If the agent's overall expected benefit from continuing the search (i.e., the reduction in price that it will obtain) is lower than the overall cost due to the added search time, the agent clearly should not continue the search. Furthermore, it has been proven that it is enough for the agent to consider only the next time period, i.e., it should stop the search if and only if the expected reduction in the price in the next time period is less than the cost of continuing one time period (λ) [31]. To understand why this is so, consider the following sketch of the proof: Suppose, on the contrary, that the agent is in time step t , the expected benefit from continuing to $t + 1$ is less than λ , but it will still be beneficial for the agent to continue until time $t' > t + 1$. The agent should then continue the search to time $t' - 1$. However, it is given that $c_{t'-1} \leq c_t$. Thus, given that the price in each time period is drawn from the same probability, the relative expected reduction of the price when moving from $t' - 1$ to t' is smaller than the expected reduction when moving from t to $t + 1$. Nonetheless, the cost per time step, λ is the same. Accordingly, we demonstrate that if it is not beneficial for the agent to continue from t to $t + 1$, it is also not beneficial to continue from $t' - 1$ to t' , contradicting our assumption that it is beneficial to the agent to continue until time period t' .

In our implementation, the prices are distributed normally with a mean μ and a standard deviation σ . We use x to denote the price for which the expected reduction in the price for one time period is equal to λ . For a given price p the benefit is $x - p$ and the probability² for p is

$$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{p-\mu}{\sigma}\right)^2}$$

Given these definitions we must generally solve:

$$\int_0^x (x-p) \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{p-\mu}{\sigma}\right)^2} dp = \lambda$$

In our specific implementation, $\mu = 1000$, $\sigma = 200$ and $\lambda = 15$. Thus we specifically solve,

$$\int_0^x (x-p) \frac{1}{200\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{p-1000}{200}\right)^2} dp = 15$$

Solving this equation yields a solution of $x = 789$.

Note that since an optimal solution exists, ostensibly there is no need for bounded rationality theories such as AAT. However, we believe that not only people, but even the agents they write on their behalf, do not necessarily effectively harness a computer's computational power to find optimal strategies. Thus, we predict that agents will use non-optimal search strategies involving instrument variables such as the current price of the commodity (x_1) and the elapsed time (x_2) as measured by the number of visited stores.

² In the domain, when a negative price was drawn, we drew a new price. Since the probability of such an event is extremely small, we did not take it into consideration in our analysis.

3.1.3 Two-Parameter Optimization Domain

In the second optimization problem, we consider an environment of a company with a monopoly for a certain product. The owners of the company must set several important variables that will impact how much money the company will make. In this environment, there are no external factors to these decisions. Thus, the outcome of these decisions is not influenced by factors such as how other companies perform, other people's decisions, or random effects.

We formally present this problem as follows: We assume that there is a finite timeline $\mathcal{T} = \{1, 2, \dots, k\}$. The agent needs to set two instrument variables, the price and quality of the product for any $t \in \mathcal{T}$, denoted p_t and q_t , respectively. The values of p_t and q_t can be set to any positive integer. The profit of the agent at a given time t depends on its choice for the price and quality until the current time. We use $\vec{p}^t = (p_1, \dots, p_t)$ and $\vec{q}^t = (q_1, \dots, q_t)$, to denote the price and qualities determined by the agent until time t . The profit of the agent at a given time, $PT(\vec{p}^t, \vec{q}^t)$ consists of the gross profit and the expenses due to the quality.

The portion of the profit that is influenced by the price is:

$$PG(\vec{p}^t) = \vec{p}_t^t e^{-(\vec{p}_t - \mu)^2 / \lambda_p^1}$$

The portion of the profit that is influenced by the quality is:

$$QG(\vec{q}^t) = \lambda_q^1 QG(t-1) + \lambda_q^2 QG(t-2) + \lambda_q^3 \sqrt{\vec{q}_t^t}$$

The profit at time t is

$$PT(\vec{c}^t, \vec{q}^t) = PG(\vec{p}^t) * QG(\vec{q}^t) - \gamma \vec{q}_t^t$$

The profit of the entire time period is: $\sum_{t \in \mathcal{T}} PT(\vec{p}^t, \vec{q}^t)$

All of the constants except μ and γ are known to the agent. In our experiments we set the constants $\lambda_p^1 = 1000$, $\lambda_q^1 = 0.7$, $\lambda_q^2 = 0.3$, and $\lambda_q^3 = 0.4$. For initialization purposes, q_{-1} and q_0 are set to 0. The value for μ is a randomly selected integer from a uniform distribution between 25 and 75, and γ is a randomly selected integer from a uniform distribution between 40 and 60. Finally, we set $k=50$ indicating that the company will only exist for 50 time periods.

The goal of the agent is to maximize the company's profit over the course of one trial. The agent operates in a one-shot environment which resets the values of μ and γ after every trial. Thus, no learning could be performed between these trials to learn the values of μ and γ . However, throughout one trial, for every time period the agent was given the values of $PT(\vec{c}^t, \vec{q}^t)$, $PG(\vec{c}^t)$, $QG(\vec{q}^t)$

Note that in this environment as well, an optimal solution can be found. Here, the only problem parameters with unknown values are μ and γ . It is possible to construct a table offline with the optimal values of price and quality given all possible permutations for μ and γ . The size of this table will be 50 (as per k) times 50 (as per all possible values for μ) times 20 (as per all possible values for γ). Once this table has been constructed, the online agent only needs to identify the values of μ and γ , so it can use the optimal prelearned values. As such, an optimal solution is as follows: In the first time period, the agent uses a predetermined value for γ that will yield the highest average profit. Once the agent observes the company's profit after the first iteration, it is able to solve for the unknown value of μ . In the second iteration it can similarly solve for γ as it is known to be the only remaining variable. After this point, the agent sets the price and quality for every remaining time step

based on the prelearned optimal values for the values of μ and γ . Alternatively, another optimal solution involves first solving for the unknown value for γ in the first iteration, for μ in the second iteration, and using the prelearned optimal values after this point.

As an optimal solution is again possible in this domain, bounded rationality theories such as AAT seem irrelevant. However, we generally believe that two significant factors contributed to the subject's inability to optimally solve these problems. First, both problems were verbally presented and thus people needed to properly model the problems before solving them. Second, even after these problems are properly modeled, correctly solving these problem parameters is far from trivial and requires significant algebraic knowledge. As a result, we again hypothesize that people will use non-optimal strategies, i.e. involving search within the instrument variables of the company's price (x_1) and quality (x_2).

While the commodity search and the monopoly optimization problems are relatively simple, the similarities and differences between them allows us to generalize our findings. First, the commodity search problem is a basic problem all people encounter – we all purchase items in stores with different price distributions. In contrast, the monopoly problem is one that business managers, and not computer science students, may have experience solving. Other technical differences exist as well. The commodity search problem is characterized by complete information, but a certain level of randomness (non-deterministic behavior) exists in what the price of the commodity is in a given store. Also note that the optimal solution involves making a decision based on the current price alone. In the first domain, other instrument variables, such as the number of visited stores or the length of the time horizon, are not part of the optimal solution. In contrast, the monopoly game is characterized by deterministic functions with two unknown parameters (μ and γ). While this problem is more straightforward, the introduction of a second variable makes the problem seemingly more complex. Furthermore, the optimal value for quality changes over time, and can only be found by solving for γ . Despite these differences, both domains have been previously found to be generalized representations of real-world problems [1,26] and thus serve as good domains for studying the models of people's search agents.

3.2 Negotiation problems

In order to study more complex, real-world problems, we also considered two challenging negotiation scenarios. As was the case within the optimization problems, we again conducted two types of experiments. First, we studied how two people negotiate with each other. The goal of this experiment was to attempt to quantify if people's decision making model is consistent with AAT or other options – whether optimal or bounded. Additionally, we studied two previously designed automated agents capable of operating in negotiation environments – the QO and KB agents. These agents represent the state of the art in automated negotiating agents, and have previously been shown to be effective in a variety of domains [11, 12, 19]. In fact, not only do these agents typically outperform the human partners they interact with, but they are even effective in significantly improving the negotiation skills of these people as well [12]. The goal of this experiment was to ascertain if these expert models followed decision making processes consistent with AAT as well.

The QO agent [11] is a domain independent agent that can negotiate with people in bilateral negotiations environments characterized by a finite horizon and incomplete information. The QO agent operates with uncertain information by applying a simple Bayesian update mechanism. Any time the agent receives an offer or a response to an offer it made, it uses this update mechanism to attempt to infer which utility best suits the opponent. The

QO agent uses a qualitative approach [28] to rank all possible offers. While the QO agent's model applies utility functions, it is based on a non-classical decision making method, rather than focusing on maximizing the expected utility. Specifically, the QO agent uses the maximin function and the qualitative valuation of offers. Using these methods, the QO agent generates offers and decides whether to accept or reject proposals it has received.

One of the main drawbacks that Lin et al. [11] state concerning their QO agent is that most of the agreements reached in negotiations in which their agent was involved were offered by the human counterpart. In addition, most of the time, their agent repeatedly proposed the same offer and did not generate new offers as time progressed.

To address this shortcoming, the KB agent was suggested as an improved variation that invoked a different strategy for the generation of offers [19]. As opposed to the QO agent, the KB agent implemented a concession oriented strategy when generating offers. They found that this agent negotiates efficiently with people and even achieved higher utility values than the QO agent in the two negotiation domains described below. Moreover, the KB agent achieved significantly better agreements, in terms of individual utility, than the QO agent and human counterparts playing the same role in these domains.

We studied two negotiation problems that are defined by different parameters that need to be negotiated. The automated agents can play either role in the negotiation, and a person accesses the negotiation interface via a web address. The simulation allows for either two human players or a human player playing against an automated agent. In either case, no programming knowledge is needed to interact with other people or agents. Instead, the negotiation session is conducted using a semi-formal language. Each player constructs an offer by choosing the different values constituting the offer. Then, the offer can be sent in plain English to the counterpart. A time effect exists that assigns a time cost which influences the utility of each player as time passes (there can be different time costs for each player). The time effect can be either negative or positive. If no agreement is reached by the end of the final turn then a status quo agreement is implemented resulting in a status quo value for each player. Each player can also quit the negotiation session at any given time if he/she decides that the negotiation session is not proceeding in a favorable way. This results in the implementation of an opt-out outcome. Finally, we assume that there are a finite set of agent types. These types are associated with different additive utility functions (e.g., one type might have a long term orientation regarding the final agreement, while the other type might have a more constrained orientation). In these negotiation scenarios, each agent is given its exact utility function. The agents, and the subjects in the experiments described in the paper, are also aware of the set of possible types of the opponent. However, the exact utility function of the rival is private information.

During each phase of the negotiation session, the instructions and parameters subject to negotiation are accessible to the players. The players are also aware of the current turn and time left until the end of the turn and until the negotiation session terminates. The history of past sessions is also easily accessible. When receiving an offer the player can choose whether to accept or reject it, or make a counter-offer.

3.2.1 Employer / Employee Negotiation Domain

Specifically, we analyzed two previously defined negotiation domains [11]. In the first problem, we consider a negotiation session that takes place after a successful job interview between an employer and a job candidate. In the negotiation session both the employer and the job candidate wish to formalize the hiring terms and conditions of the applicant. Below are the issues under negotiation:

1. **Salary.** This issue dictates the total net salary the applicant will receive per month. The possible values are (a) \$7,000, (b) \$12,000, or (c) \$20,000. Thus, a total of 3 possible values are allowed for this issue.
2. **Job description.** This issue describes the job description and responsibilities given to the job applicant. The job description has an effect on the advancement of the candidate in his/her work place and his/her prestige. The possible values are (a) QA, (b) Programmer, (c) Team Manager, or (d) Project Manager. Thus, a total of 4 possible values are allowed for this issue.
3. **Car benefits.** In addition to the base salary, other job benefits may also be negotiated. This issue revolves around the possibility that the company will provide a company car for use by the employee. Possible values for this issue are (a) providing a leased company car, (b) no leased car, or (c) no agreement.
4. **Pension benefits.** Additionally, we assume that pension benefits must also be negotiated. The possible value for the percentage of the salary deposited in pension funds are (a) 0%, (b) 10%, (c) 20%, or (d) no agreement. Thus, a total of 12 possible values ($3 \times 4 = 12$) are allowed for this issue.
5. **Promotion possibilities.** This issue describes the commitment by the employer regarding the fast track for promotion for the job candidate. The possible values are (a) fast promotion track (2 years), (b) slow promotion track (4 years), or (c) no agreement. Thus, a total of 3 possible values are allowed for this issue.
6. **Working hours.** This issue describes the number of working hours required by the employee per day (not including over-time). The possible values are (a) 8 hours, (b) 9 hours, or (c) 10 hours. Thus, a total of 3 possible values are allowed for this issue.

In total, this scenario allows for a total of 1,296 possible agreements ($3 \times 4 \times 12 \times 3 \times 3 = 1296$).

Each turn in the negotiation simulation is equivalent to two minutes of an actual negotiation session, and the total negotiation session is limited to 28 minutes. If the sides do not reach an agreement by the end of the allocated time, the job interview ends with the candidate being hired with a standard contract, which cannot be renegotiated during the first year. This outcome is modeled for both agents as the status quo outcome.

During negotiation, each side can also opt-out of the negotiation session if it feels that the prospects of reaching an agreement with the opponent are slim or if they feel it is no longer able to negotiate a fair deal. If the employer opts out then she will incur an expense due to the lost time and work from the potential employee. As we assume the employer will be required to postpone the project for which the candidate was interviewing, this cost can be considerable. Conversely, the employee also has some leverage. If the employee accepts too little, he is likely to find better work elsewhere. However, the employee also loses the time and salary from lost wages in continuing their job search. Additionally, opting-out will make it very difficult for him to find another job, as the employer will spread his/her negative impression of the candidate to other CEOs of large companies.

Time also has an impact on the interaction. As time advances, the candidate's utility decreases as the employer's good impression has of the job candidate decreases. The employer's utility also decreases as the candidate becomes less motivated to work for the company.

3.2.2 Political Dispute Negotiation Domain

In this problem England and Zimbabwe negotiate in order to reach an agreement evolving from the World Health Organizations Framework Convention on Tobacco Control, the

world's first public health treaty. The principal goal of the convention is "to protect present and future generations from the devastating health, social, environmental and economic consequences of tobacco consumption and exposure to tobacco smoke." The leaders of both countries are about to meet at a long scheduled summit and must reach an agreement on the following issues:

1. Creation of a Global Tobacco Fund. This issue describes the total amount to be deposited into the Global Tobacco Fund to aid countries seeking to rid themselves of economic dependence on tobacco production. This issue has an impact on the budget of England and on the effectiveness of short-range and long-range economic benefits for Zimbabwe. The possible values are (a) \$10 billion, (b) \$50 billion, (c) \$100 billion, or (d) no agreement. Thus, a total of 4 possible values are allowed for this issue.
2. Impact on other aid programs. This issue affects the net cost to England and the overall benefit for Zimbabwe. If other aid programs are reduced, the economic difficulties for Zimbabwe will increase. The possible values are (a) no reduction, (b) reduction equal to half of the Global Tobacco Fund, (c) reduction equal to the size of the Global Tobacco Fund, or (d) no agreement. Thus, a total of 4 possible values are allowed for this issue.
3. Trade issues – Zimbabwe. Both countries can use trade policy to extract concessions or provide incentives to the other party. They can use restrictive trade barriers such as tariffs (taxes on imports from the other country) or they can liberalize their trade policy by increasing imports from the other party. There are both benefits and costs involved in these policies: tariffs may increase revenue in the short run but lead to higher prices for consumers and possible retaliation by affected countries over the long run. Increasing imports can cause problems for domestic industries. But it can also lead to lower consumer costs and improved welfare. For Zimbabwe, possible stances are: (a) reducing tariffs on imports, (b) increasing tariffs on imports, (c) no change.
4. Trade issues – England. Similarly, England must decide on the same trade issues to: (a) reduce, (b) increase, or (c) leave tariffs unchanged. Taking both sides into consideration, a total of 9 possible values are allowed for both sides of these issues.
5. Forum to Study Long-Term Health Issues. This issue revolves around the scope of a forum to explore comparable arrangements for other long-term health issues. This issue relates to the precedent that may be set by the Global Tobacco Fund. If the fund is established, Zimbabwe will be highly motivated to apply the same approach to other global health agreements. This would be very costly to England. The possible values are (a) creation of a fund, (b) creation of a committee to discuss the creation of a fund, (c) creation of a committee to develop an agenda for future discussions, or (d) no agreement.

Consequently, a total of 576 possible agreements exist ($4 \times 4 \times 3 \times 3 \times 4 = 576$). While for the first two issues there are contradicting preferences for England and Zimbabwe, for the last two issues there are options which might be jointly preferred by both sides.

Each turn in the simulation is equivalent to a week of the summit, while the summit is limited to 14 weeks. If no agreement is reached within the specified time limit, the Framework Convention will be seen as an empty document, devoid of any political significance. This will be a blow to England, which has invested political capital to reach an agreement, in the hope of gaining support for other, perhaps more important, international agreements in the future. It will also, however, save England money in the near term. For Zimbabwe, failure to reach an agreement will create a major financial hardship and deprive it of a precedent that can be used for future negotiations. Consequently, England is better able to accept a failure than is Zimbabwe. This outcome is modeled for both agents as the status quo outcome.

Opting out of the negotiation is also an option. Opting out by England means trade sanctions imposed by England on Zimbabwe (including a ban on the import of tobacco from Zimbabwe), while if Zimbabwe opts out then it will boycott all British imports. However, if England opts out it also saves the funds that would have been spent on the Tobacco Fund, and if Zimbabwe opts out it loses the opportunity for financial gain and for assistance in reducing the health problems that arise from tobacco use. Consequently, England will likely be more willing to opt out if the negotiations are not going its way, and Zimbabwe will be more willing to continue negotiations until an agreement is reached.

Time also has an impact on the negotiations. Creation of the fund is more urgent for Zimbabwe than for England. Consequently, Zimbabwe has an incentive to reach an agreement earlier rather than later; thus as time advances Zimbabwe's utility reduces. On the other hand, England gains as time advances, as it postpones the time at which it must transfer money to the fund.

Taking into account the different types of agents, we can say, for example, that an agent representing Zimbabwe with a short term orientation will focus on a short term redistribution of resources, insist on the largest possible current assistance and help with long-term health problems, as well as trade concessions. On the other hand, an agent representing England with the same short term orientation, for example, will aim to minimize current cost, limit impact on trade, and maintain its economic and political position in the near term.

4 Results and Analysis

To the best of our knowledge, this paper represents the first of its kind to demonstrate that many agents designed to solve optimizing and negotiation problems in fact implemented strategies consistent with key elements of AAT. In this section we present two basic types of experiments to support this claim – those where people's strategies were studied, and those analyzing computer agent's strategies. In Section 4.1 we study 70 people's strategies to solve the two previously described optimization problems. As previously described, these strategies were written in plain text without any need to program them. In Section 4.2 we study nearly 100 agents programmed for the same problems. We demonstrate that the performance of these agents is far from optimal. In both Sections 4.1 and 4.2, we base the analysis of these people's and agent's decision models on 5 judges who overwhelmingly classify them as being more consistent with AAT than traditional rational models or other bounded theories. We also tested for statistical significance in the differences between judges through one-way ANOVA tests and found them to be statistically insignificant in all experiments. Thus, we can confidently claim that the results in these sections confirm that people, and the agents they write to operate on their behalf, both adhere to AAT within optimization problems. In Section 4.3, we find that key elements of AAT also exist in the peoples' and agents' strategies in the negotiation problems. To support this claim, we present results from a total of over 100 negotiation sessions between people (human negotiation with human) as well as 33 sessions between the KB and QO agents with people in each of the negotiation domains (for a total of $33 \times 2 \times 2$ or 132 negotiation sessions between agents and people).

4.1 Analysis of People's Optimizing Strategies

In order to study people's strategies in the two optimization problems, we conducted the following experiment. First, we present each person with a web interface simulating the

problem at hand³. Next, we allowed each person to interact with the simulation until they felt that they could describe a general strategy for how they would solve the problem. Finally, we asked each person to describe their strategy in plain text. Note that as most of these people had minimal to no programming experience, no code or pseudo-code was used to describe these strategies.

We found that these strategies generally exhibit three key elements specific to AAT. First, we discovered that people prioritized certain instrument variables to be solved first. This concept parallels AAT’s concept of urgency between its m goal variables. Second, we found that people generally stopped their search within a given instrument variable once it had satisfied this value. Finally, we found that people often changed the satisficing threshold during the search process. This parallels AAT’s concept of retreat and should be expected if the people deem its original goal is no longer realistic. For example, within the commodity search problem, people described strategies where they first attempted to buy the commodity below a certain threshold price. However, assuming a certain time elapses (or a number of stores have been visited), they revise downward this threshold as being a “good enough” solution. In the monopoly domain, they often prioritized a given variable, typically the price variable, to be searched for first. After this variable was satisfied, they proceeded to the other variable, typically the quality variable, either with a second search procedure or according to some fixed value.

It is important to note that two key differences exist between classic AAT, and the behavior exhibited in the strategies we studied. First, AAT assumes that the m goal variables used to solve \mathcal{G} are incomparable since no utility function is possible to connect goal variables. Here, we consider problems where a concrete function between \mathcal{G} and the m goal variables clearly exists. Nonetheless, we hypothesize that people will not attempt to calculate \mathcal{G} due to its bounded nature. As we describe in Section 2, this represents a significant generalization to AAT’s theory and its relevance even in domains that contain concrete, albeit difficult to quantify, utility functions. Second, AAT is based on the premise that the person’s search will be based on an *aspiration scale* which sorts the m goal variables and attempts to satisfice values for these goals. As we consider concrete optimization problems, it is more natural for people to consider optimizing the instrument variables that constitute the basis of these goals rather than the more abstract general goal variables. For example, we would expect a person in the monopoly domain to focus on variables such as price and quality instead of higher level abstract goals such as “brand awareness” or “company infrastructure”. Again, we believe that this difference is due to the fact that we consider optimization problems with clear utility functions, and this difference again represents a generalization of AAT.

	Judge 1	Judge 2	Judge 3	Judge 4	Judge 5	Average	AAT?
Price	3	5	10	14	9	8.2	No
Stores Visited	13	12	11	10	12	11.6	No
Both w/ Retreat	29	28	23	28	30	27.6	Yes
Predefined	4	0	0	0	0	0.8	No
Other	4	8	9	1	2	4.8	No

Table 1 Goal Variables in the Commodity Search Domain without Programming

³ For example, the directions and web interface for the one parameter commodity search problem can be seen at <http://www.jct.ac.il/~rosenfa/costSearch.html>.

The five judges found that most people described strategies consistent with AAT. Table 1 displays the analysis of 53 participants' strategies for the one parameter, commodity search problem. This table depicts how the judges categorized the number and search variables of the subjects. The optimal solution within this problem involves search within only one variable, the current price of the commodity. Nonetheless, the judges found that on average 8.2 of the strategies made decisions based on this variable alone (see row 1, column 6). While none of these subjects actually described the optimal strategy (buy if the price in the current store is less than 789 - a solution dictated by previous research [31]) several subjects did have similar strategies such as buy if the price is less than 800. An average of 11.6 of the strategies actually used another search variable, the number of stores visited, to make decisions (see row 2). For these subjects, the strategy was to visit a predetermined number of Y stores and buy the commodity in the cheapest store from the group of Y . Both of these strategies only contained one instrument variable. As such, we can view them as fast and frugal search strategies [2] involving only one variable (e.g. search until price $< X$, or visit Y stores and buy in the cheapest store). Only one judge believed a small number of subjects used simple fast and frugal predefined strategies (row 4). Note that here, some of the strategies could not be clearly identified due to ambiguities in the text strategies and were left as being categorized as Other (line 5).

As is clear from these results, even without the need to program their strategies, the majority of people (average 27.6 of 53) still choose a strategy based on AAT (line 3 – Both w/ Retreat). For these subjects, the strategy was to immediately buy the commodity if the price was less than X , otherwise, they visited a maximum of Y stores and bought in the store with the cheapest found price. While non-optimal, this strategy has key elements of AAT. Originally, people searched based on price alone. However, if the desired price was not found, the subject revised his/her aspiration downward. This can be seen as being similar to the retreat process in AAT's goal variables. Note that the use of *urgency* here is not at all justified based on optimal behavior as the second goal variable (Stores Visited) is not even part of the optimal solution. Furthermore, the combination strategy of settling on a price after visiting Y stores if no commodity is found with a price less than X is a good example of *retreat* values. Here, a price less than X was aspired for. However, assuming this could not be found after visiting Y stores, the satisficing threshold of this variable was revised and a lower value was accepted. Thus, these results demonstrate that people often use concepts of bounded rationality even in domains with clear utilities.

Case	Parameter1	Parameter2	Judge1	Judge2	Judge3	Judge 4	Judge 5	Average	AAT?
1	Optimal	Optimal	0	0	0	0	0	0	No
2	Simult. Search	Simult. Search	3	3	5	6	6	4.6	No
3	Predetermined	Predetermined	1	1	0	0	0	0.4	No
4	Search Price	Search Quality	3	5	6	7	6	5.4	Yes
5	Search Quality	Search Price	1	1	1	3	1	1.4	Yes
6	Search Price	Predetermined Quality	12	10	6	4	7	7.8	Yes
7	Search Quality	Predetermined Price	0	0	1	0	0	0.2	Yes
8	Alternating	Alternating	0	0	1	0	0	0.2	Yes

Table 2 People's Strategies in the Monopoly (Price and Quality) Domain without Programming

Similarly, we studied a total of 20 subjects' strategies used to solve the two parameter monopoly optimization problem. Most subjects (lines 4 & 6) searched for the price variable

first (e.g. this parameter was given the highest urgency) and afterwards quality was either searched for (in line 4) or set by a predetermined function (in line 6). A very small number of people had the opposite *aspiration scale* and first searched for quality (in lines 5 & 7). Similarly, only a relatively small number of people (line 8) consistently alternated between searching for price and quality. In this approach, a person would set one value (say price), in the next time frame search for the optimal value of the second goal (quality), only to return back to the first goal variable (price) and adjust downward (retreat from) its original value. In either case, the vast majority of strategies (average 15 of 20) were deemed to be consistent with AAT.

4.2 Analysis of Optimizing Agents

As per the second part of our research methodology, we also asked a second group of subjects to program their strategies. This allowed us to analyze the agent’s strategies and performance. Overall, in both optimizing domains, a minority of the agents did in fact exhibit performance near or close to that of the optimal agent. However, the vast majority of the agents, regardless of the programming background of the subjects, deviated significantly from optimal behavior. We then explored the decision models used by the sub-optimal agents and found them to also be consistent with AAT.

In accordance with previous work by Chalamish et al. [1], 41 agents from the commodity search domain were divided into 14 “maximizing” agents and 27 “cloning” agents. The “maximizing” agents were written by subjects who were asked to create agents with as high performance as possible (optimal). The “cloning” agents were written by subjects who were instructed to mimic their own personal strategies. Since one focus of the experiment was how effectively people could clone their own strategies, the maximize group served as the control group with a ratio of 1:2. As random effects do exist in this environment, we ran each of these agents 50 times, and averaged the agent’s performance. We then compared the average performance from the “cloning” and “maximizing” groups, the best performing agent from each of these groups, and the worst performing agent from each of these groups. Finally, we compared the performance of the optimal agent to that of the agents the subjects wrote.

Figure 1 depicts the performance of these agents. Note that the goal of this domain was to minimize the search price. As such, the low average search cost of the best agent (column 3) closely approximated the performance of the optimal agent (column 4). It is important to also note that the average cost of both the “cloning” and “maximizing” agents (approximately 830 units) were quite far from the optimal agent (approximately 790) with p-values testing for significance being much less than 0.0001. However, the differences between the “cloning” and “maximizing” agents were not significant (p-value 0.48). These results imply that most people asked to write optimal agents fall well short of this amount. Instead, subjects, whether they are part of “cloning” or “maximizing” groups, seem to implement similar, non-optimal strategies.

We then studied a second group of 57 maximizing agents written for the monopoly domain. We recorded the average, highest, and lowest performance across the agents, and compared this performance to that of the optimal agent. In this domain, many different values were possible for the previously described price and quality functions. As such, we applied two different evaluation approaches. In the *Full* evaluation, we studied the average performance of the agents across all possible permutations of price and quality. In the *Sampling* evaluation, we studied the average performance of the agents across six randomly selected

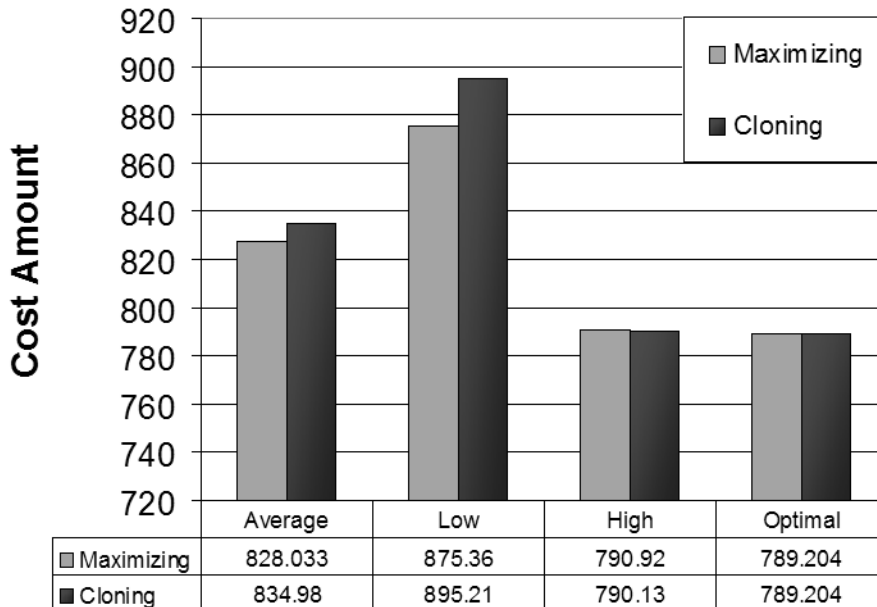


Fig. 1 Comparison of the Average, Worst, and Best Cost Amount of “Maximizing” and “Cloning” Commodity Search Agents to Optimal Values. Lower Costs are Better.

value pairs of these values. Realistically, the second type of evaluation seems more appropriate as people typically build small numbers of companies. However, the fuller evaluation in the *Full* group is useful for statistical testing.

Figure 2 displays the performance of the agents from the monopoly domain. Note that in this domain as well, the performance of the best agents (third column) in both the *Full* and *Sampling* evaluation methods reached near optimal levels. However, the agents’ average performance (first column) again fell well short of the optimal (p-values between the optimal performance and the Full and Sampling evaluation groups were both well below 0.0001). This again strongly supports the claim that people’s agents typically fall far short of optimal values. Note that in the more complex monopoly domain, the average agent’s performance was over 30% less than the optimal value, while in the simpler commodity search problem, this difference was closer to only 5%. This seems to imply that as problems become progressively harder, the agents people write to perform in these problems become more bounded and deviate further from the optimal values.

Given that these agents do not use optimal strategies, we wanted to generalize what non-optimal strategies were in fact being used. Table 3 presents the analysis of the 5 judges of the programmed agents. These results can be compared to the results presented in Table 1. Note that for these agents an even larger majority, or over 71% of the agents (see row 3 column 6, average 29.2 of 41) used combination strategies consistent with key elements of AAT.

Similarly, the results in Table 4 represent the analysis of the five judges of the monopoly agents’ decision processes. None of the agents used the optimal strategy of solving for both the price and quality variables (line 1). A number of agents did simultaneously search for

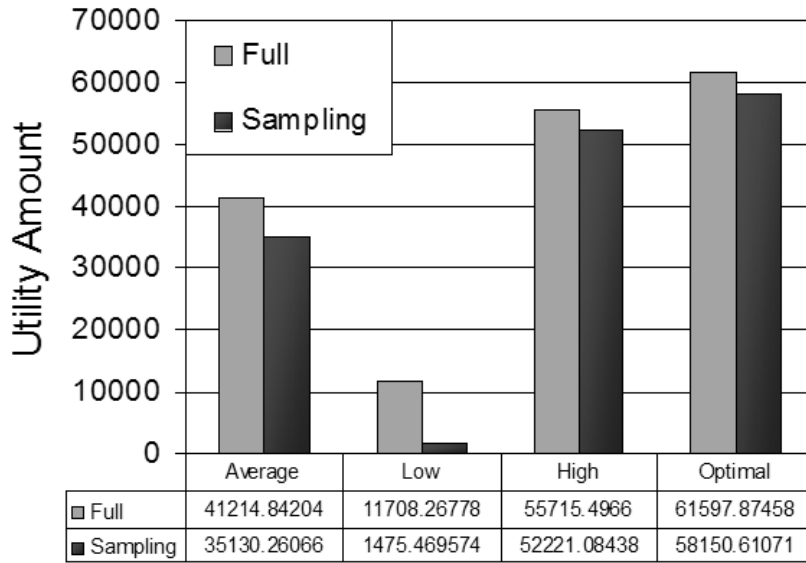


Fig. 2 Comparison of the Average, Worst, and Best Utility Value of Monopoly Agents to Optimal Values. Higher utilities are better.

	Judge 1	Judge 2	Judge 3	Judge 4	Judge 5	Average	AAT?
Price	6	6	7	9	4	6.4	No
Stores Visited	5	5	5	7	5	5.4	No
Both w/ Retreat	30	30	29	25	32	29.2	Yes

Table 3 Goal Variables in the Commodity Search Domain with Programming

both the price and quality variables (line 2), and a smaller number of subjects did use predetermined heuristics for setting both variables (e.g. price always equals 10 and quality equals time elapsed). Both of these strategies do not contain elements of AAT as no urgency exists between variables (in line 2), or no search for goal variables was performed (in the predefined case). These agents must be categorized as being consistent with other methods – such as more classic simultaneous search approaches [21] (in line 2) or fast and frugal heuristics [2] (in line 3).

Again, for the vast majority of the agents (over 73%) search was conducted with elements of AAT. Most agents (lines 4 & 6) searched for the price variable first (e.g. this variable had the highest urgency) and afterwards quality was either searched for (in line 4) or set by a predetermined function (in line 6). A very small number of agents had the opposite *aspiration scale* with quality searched for first (in lines 5 & 7). Similarly, only a relatively small number of agents (line 8) consistently alternated between searching for price and quality. In this approach, an agent would set one value (say price), in the next time frame it would search for the optimal value of the second goal (quality), only to return back to the first goal variable (price) and adjust downward (retreat from) its original value. The reason

Case	Parameter1	Parameter2	Judge 1	Judge 2	Judge 3	Judge 4	Judge 5	Average	AAT?
1	Optimal	Optimal	0	0	0	0	0	0	No
2	Simult. Search	Simult. Search	9	13	8	13	11	10.8	No
3	Predetermined	Predetermined	3	3	4	9	3	4.4	No
4	Search Price	Search Quality	9	14	23	12	20	15.6	Yes
5	Search Quality	Search Price	4	0	1	3	4	2.4	Yes
6	Search Price	Predetermined Quality	27	22	15	14	7	17	Yes
7	Search Quality	Predetermined Price	0	1	0	2	9	2.4	Yes
8	Alternating	Alternating	5	4	6	4	3	4.4	Yes

Table 4 Agent’s Strategies in the Monopoly (Price and Quality) Domain with Programming

fewer agents made use of retreat in this domain is evidently as follows. According to AAT, retreat occurs once an agent realizes it must change its aspiration based on the infeasibility of satisfying multiple goals. In the monopoly game subjects typically did not see any infeasibility element and therefore did not retreat from the values they had set. To confirm this theory, we reviewed the questionnaires the subjects filled out after describing their strategies. In this domain, many subjects described that they had solved for, or found the best price value, and thus had no need to retreat or revisit this parameter. However, in the commodity search problem, subjects (albeit wrongly) perceived infeasibility and therefore retreated from their values. Thus they made use of *retreat* variables to refine their aspirations.

4.3 Negotiation Results and AAT

In the negotiation domain, we found that people, and the expert QO and KB negotiation agents designed to interact with them, contained several key elements of AAT. Both the people negotiating among themselves and the expert agents typically lowered the value they proposed for one parameter and raised the value they desired for a different value. When this occurs, it is a classic instance of AAT’s concepts of urgency and retreat as the decision maker clearly aspires for a higher value for one variable (as defined by its urgency), but is willing to settle on a lower value (retreat) for a different variable in order to achieve this goal. Note that this type of negotiation is different from approaches that change several parameters simultaneously, as can be done in traditional machine learning. It is also different from approaches that use simple heuristics to simplify the decision process. It is important to note that AAT theory was originally formulated for domains where no utility could be computed, and thus AAT generally refers to goal variables and not to urgency and retreat of the specific values of variables. In contrast, we found AAT was even applicable where these utilities did exist, again constituting a significant generalization of AAT in this domain as well.

Furthermore, it is important to stress that both the KB and QO expert automated negotiation agents were intentionally designed **not** to use optimal equilibrium strategies that address multiple parameters simultaneously. Previous work found that in order to successfully interact with people, automated negotiation agents must be designed to deviate from optimal strategies. This is because they found that when automated agents follow an equilibrium strategy the human negotiators negotiating with them become frustrated. This is mainly due to the fact that automated agents repeatedly propose the same offer, and negotiation sessions often end with no agreement [9]. Similarly, Lin et al. [11], created an optimal agent based on the equilibrium policy for the above described negotiation problems. They also found that in many situations the equilibrium agent fails to reach any agreement when

playing with people [11]. Thus, both the KB and QO agents also intentionally steered away from a design based on theoretically optimal equilibrium strategies [11]. In fact, a survey of articles on automated negotiators [10] found that none of the reviewed agents implemented equilibrium strategies. They attributed this result to the assumption that optimal strategies fare poorly when people must negotiate in a complex environment due to their bounded nature.

As in the two optimization problems, we also studied if and to what degree people's negotiation agreements deviated from the equilibrium strategies in these two problems. To study this point, we examined the utilities yielded by the previously studied equilibrium negotiation strategies for these problems [11]. However, there are two options for the equilibrium protocol depending on which side is allowed to make the last offer. As a result, we computed the average between the expected utility of both possibilities, and then checked the proximity of the utility of people's actual agreements to this amount. We found that in the employer / employee domain people's agreements on average yielded a utility of 452.83 versus the equilibrium utility of 569.43 (over 20% worse than the equilibrium policy). In the Tobacco Trade Domain this difference was even more significant, with the average agreement between people yielding a utility of only 326.61 versus 816.93 of the equilibrium policy (60% worse than the equilibrium policy). Thus, in both cases, one must ask if bounded models, such as AAT provide a better model for interactions in this domain.

Based on AAT, we expected to find two phenomena in these results. First, in accordance with AAT, agents and people would generally focus on raising and lowering only one parameter at a time. Thus, negotiations would typically not involve many parameters simultaneously. Second, we would expect that certain parameters would be more important than others. In other words, the aspiration scale would typically be similar across different people, albeit with some possible fluctuations. For example, in the employer / employee job negotiation domain certain people may wish to focus their negotiation on salary or work hours. Other parameters, such as long-term promotion possibilities, may be secondary concerns and not as stressed in the negotiation.

Table 5 provides support that people, and the expert agents programmed to interact with them, generally focused on only one parameter at a time. In this table, we present the average of parameters raised (left side of Table 5) and lowered (right side of Table 5) in both the Employer / Employee negotiation domain (line 1) and the Tobacco negotiation domain (line 2). For the results from the QO and KB agents (the first two columns of Table 5), we studied a total of 33 sessions between the agents and people. The human results (the third column in Table 5) represent data from 47 people interacting with other people in the Employer / Employee negotiation domain and 56 people in the Tobacco Trade domain. Note that people in both domains typically raise and lower slightly more than one parameter per negotiation turn (between 1.29 and 1.38 across all values). Both the QO and KB agents similarly do not typically raise or lower most parameters for any given negotiation turn (on average between 1.16 and 1.55 parameters across both agents in both domains). We studied the statistical similarity between the parameters raised and lowered in both domains by the automated agents versus those of people. Each statistical test contained 44 data points as both the agents and the people could change a total of 22 parameters consisting of raising or lowering 6 work parameters and 5 trade parameters ($2 \times 6 + 2 \times 5$). Both the QO and KB agents clearly failed the t-test testing for significant deviation, showing that they did not deviate significantly from the average aspiration scales of people. However, the p-score for the QO agent was much higher than that of the KB agent. Similarly, the statistical correlation between the QO agent and the parameters changed by people was much higher than that of the KB agent (see line 3 of Table 4). This result implies that the offers generated by the QO agent more

similarly mimics those made by people. Overall, both people and the expert agents did not typically raise and lower multiple parameters, showing that complex agreements involving all parameters were not pursued.

	QO	KB	Human		QO	KB	Human
Work Average Up	1.31	1.24	1.38	Average Down	1.42	1.2	1.37
Trade Average Up	1.17	1.16	1.33	Average Down	1.55	1.22	1.29
t-Test Score (p-score)	0.97	0.42	-	Statistical Correlation	0.53	0.13	-

Table 5 Comparison of the average number of parameters that were revised upward (left side) and lowered (right side) in the Employer / Employee negotiation domain (first line) and in the Tobacco Trade Domain (second line).

In order to study the second point regarding specifics of the aspiration scales, we compiled data about parameters which were typically raised or lowered. Figures 3 and 4 present these findings from the work negotiation domain and trade negotiation domain, respectively. In both Figures, we present the percentage of times each of the parameters were raised (top half of Figures 3 and 4) and lowered (bottom half of Figures 3 and 4) over all negotiation sessions. Each data point in these Figures represents the percentage of times a given parameter was raised or lowered from the total number of offers. For example, in the work domain, the 47 negotiation sessions between people yielded a total of 234 offers that revised parameters upwards. Of these, only 28 offers (12%) included the promotion possibility parameter. In contrast, the job description parameter was raised higher in 61 (26%) of the 234 offers.

As a baseline, we also considered the actual weights these issues were given [11]. This allowed us to compare if the aspiration scales people and / or the automated agents used were consistently influenced by the actual weight in the utility function for the negotiation problems. In general, we found that the aspiration scales were not necessarily consistent with the actual utility functions. For example, for the Tobacco Trade problem between people the actual weight of the Size of Fund issue was 50%, with all of the other issues combined comprising the remaining 50%. Nonetheless, note in Figure 4, that people typically addressed other issues with greater frequency, especially the Impact of Aid and the Forum on Other Health issues. Thus, clearly the actual weights of these issues' utility cannot serve as an accurate heuristic for estimating people's or agent's aspiration scales.

Several additional items are noteworthy in reference to these results. First, as is consistent with the data presented in Table 5, people and agents typically do not raise or lower all given parameters. In other words, the average sum of all parameters raised or lowered in any given offer is close to one as predicted by AAT. However, also note two findings about the differences in urgency and aspiration scales between agents and people. As per AAT theory, urgency refers to the parameter which is the most important to satisfy. Agents and people often have a clear preference for raising or lowering one parameter. For example, note that in the top half of Figure 4, the agents and the people often revised the "Forum on Other Health Issues" parameter upward in their offers, but less frequently (and never with the KB and QO agents) revised the "Size of Fund" parameter upward. Similarly, note in the top of Figure 3 that the QO agent most frequently raised the "Job Description" parameter, but never raised the "Leased Car" parameter. Also, notice that the aspiration scales which dictate the parameters typically raised and lowered also differ across agents. Again, note in Figure 3 that the QO agent upwardly raised the "Job Description" parameter much more than the KB agent and human subjects, but the "Working Hours" parameter, which had the highest urgency among human participants, was largely ignored by the automated agents.

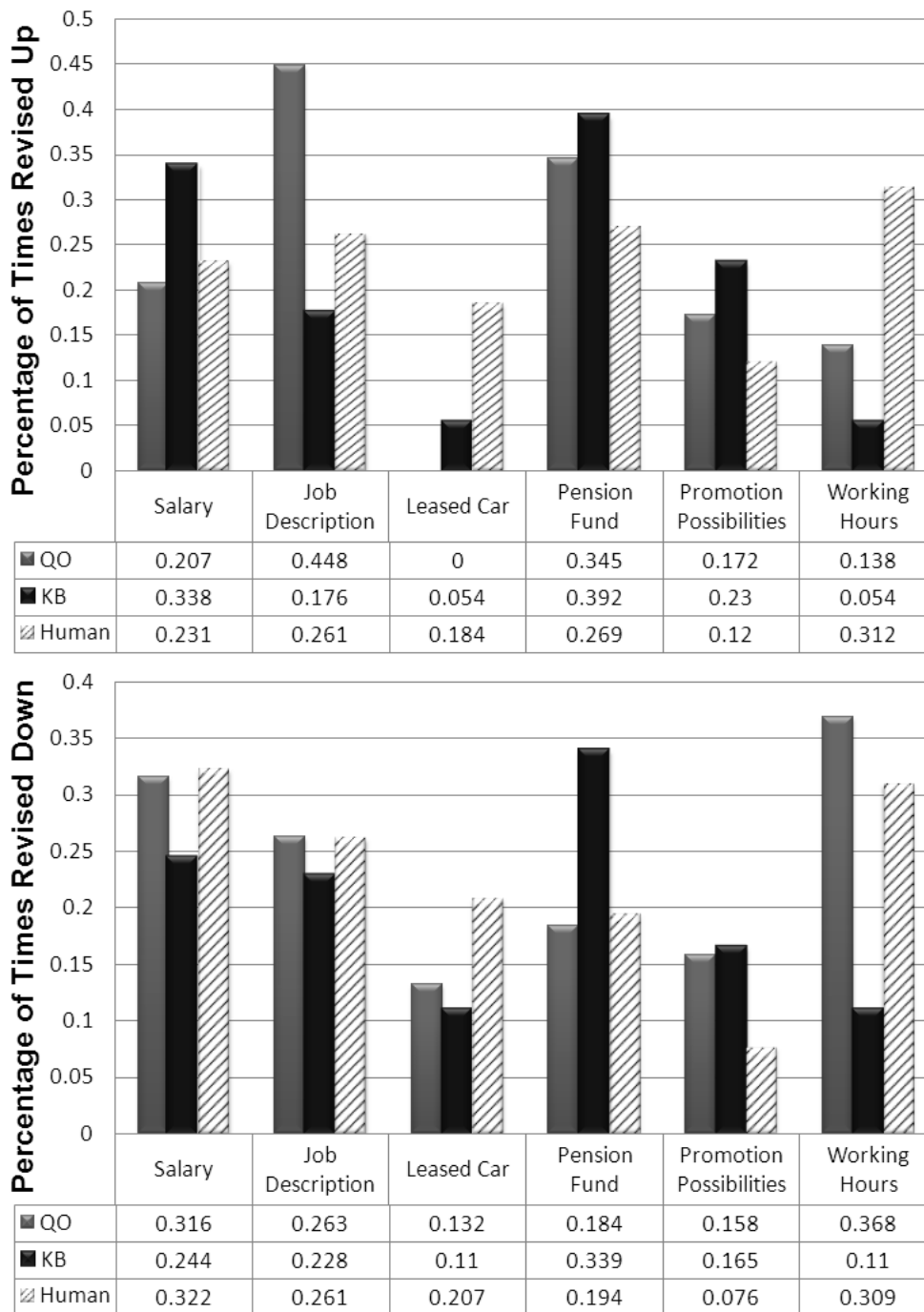


Fig. 3 The graphs present the percentage of times each of the 6 parameters was revised up (top portion of graph) and down (lower portion of graph) in the employer / employee job negotiation. The two left bars represent the parameters changed by the QO and KB agents in negotiation sessions with people, and the right bar represents the parameters changed in offers exclusively between people.

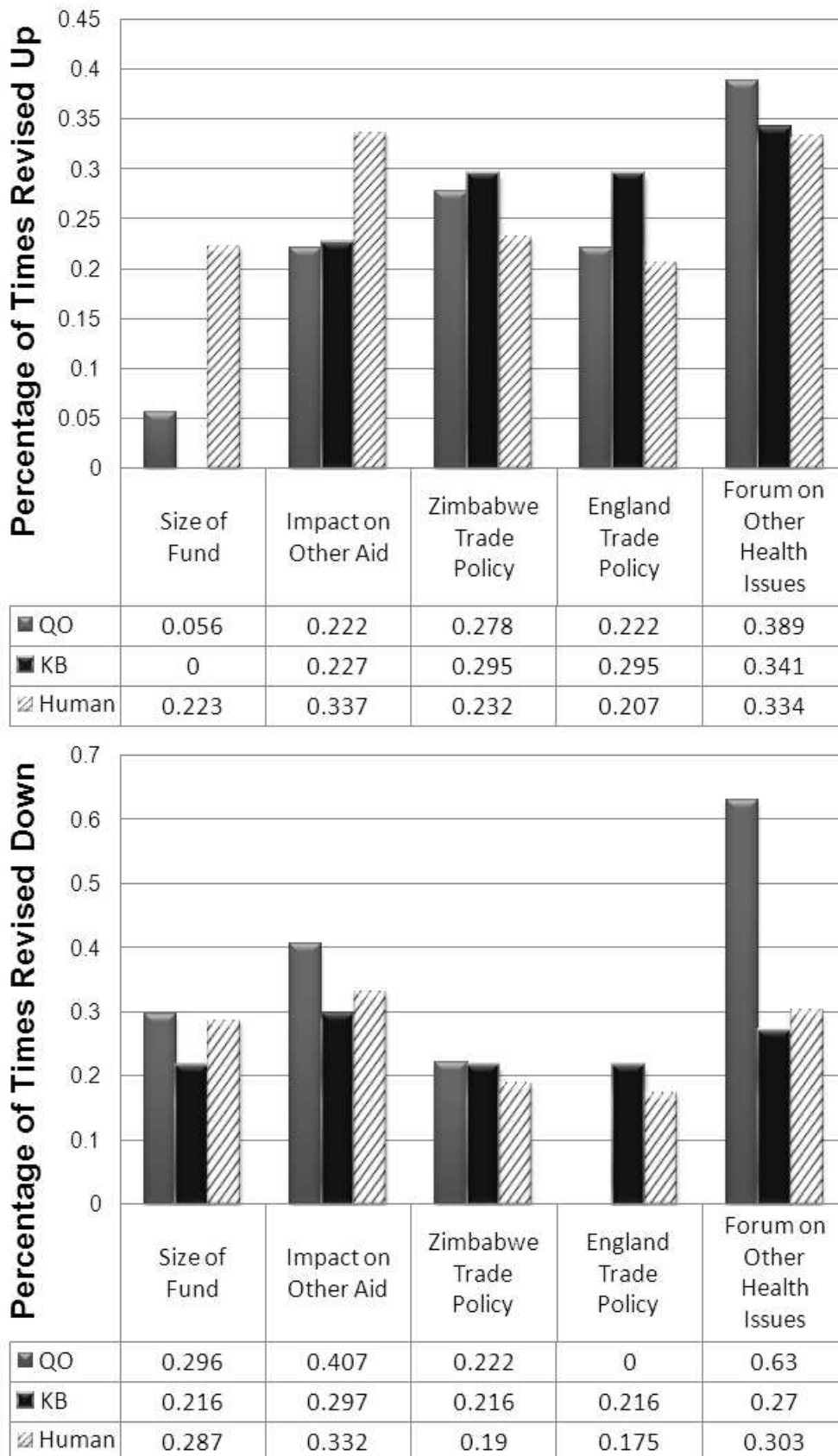


Fig. 4 The graphs present the percentage of times each of the 5 parameters were revised up (top portion of graph) and down (lower portion of graph) in the Tobacco Trade negotiation domain. The left two bars represent the performance of the QO and KB agents and the right bar represents the parameters changed in offers between people.

Thus, while both of these agents are based on mathematical models that did not explicitly take into consideration bounded models such as AAT, we did find that both these agents nonetheless did typically act consistently with AAT. However, the exact urgency and aspiration scales used by human participants and automated agents were not necessarily the same. In the next section we discuss how future agents might be designed to explicitly draw upon the lessons learned from AAT and thus perform even better.

5 Discussion and Future Applications

Overall, several conclusions can be drawn from the results of the two problem sets in both the optimizing and negotiation problems. In the optimizing problems, nearly all strategies, both of the people and their agents, were consistent with key elements from AAT. Referring back to Tables 1-4, note that typically over 70% of the strategies were classified by the 5 judges as being consistent with AAT. Furthermore, the findings presented in this paper confirm the validity of the strategy method, since the strategies people described they would use in the optimization domains, without any need to program them, were still quite similar to the strategies a different group of people implemented in their agents. Only a minority of the strategies used other models such as classic simultaneous search approaches traditionally used in artificial intelligence [21] or alternate bounded rationality models such as fast and frugal heuristics [2]. Thus, we conclude that optimal approaches cannot properly model most people's agents, and that AAT is the best alternative to be used instead. Similarly in the negotiation domain problems, both automated negotiators were not programmed with optimal equilibrium strategies or fast and frugal heuristics. Instead they contained elements of AAT as each negotiation session involved an aspiration scheme [23] where one parameter was adjusted upward and a second parameter retreated from its value.

Two possible variations of the described research methodology provide interesting directions for future work. First, the research methodology described in this paper confirms the existence of elements of AAT in the two optimization problems, as verified by the human judges. An interesting variation would be to restrict all analysis of people's performance to the log files of their decisions. We could then consider whether traditional machine learning algorithms can learn exactly which strategies are being used, and thus remove any possible human bias from the results in this paper. A second interesting variation would be to check whether the strategies which were consistent with AAT performed better than those which did not include elements of AAT. We hope to further study whether or not a link can be found between AAT based strategies and performance.

We also believe that these findings can help people create more successful agents that better simulate human performance or interact with people. We propose that the following lessons be learned from the optimizing problems: First, in simulated human behavior, even in moderately complex optimization problems, optimal methods should not be used as they do not realistically encapsulate most humans' behavior. Instead, bounded methods should be created such as those based on AAT. In understanding the strategies used by people, we propose that a small pilot be used based on the strategy method [25]. This should identify the ordering (urgency) for search variables and a range of aspiration values within these variables. For example, in the domains we studied, such a pilot would clearly identify price as the variable first searched for in both domains. Finally, any pilot should be focused on the range of aspired for values for each goal variable such that some distribution can be constructed to realistically model the range of problem solving approaches. In this way one can better create realistic agents that simulate human behavior in these and similar domains.

In reference to the negotiation problems, we believe that explicitly incorporating the lessons learned from AAT can be helpful in creating a new generation of more successful automated agents for negotiating with people. We noted in the previous section that while both the QO and KB agents had key elements of AAT, the QO agent's negotiation offers seemed to be more similar to those of people. Thus, one might assume that the QO agent is more successful in interacting with people. However, just the opposite is true. As has been previously found [19], the QO agent's weakness is its lack of a strong compromise mechanism. Thus, it does not always effectively counter offer in negotiation settings, leading to less effective agreements.

To illustrate this point, we studied the average number of offers made by the QO and KB agents compared to the number of offers made by people. In the Employer / Employee domain these results are presented in the fourth row of Table 5 and for the Tobacco Trade domain they are given in sixth row of Table 5. Note that in both domains people typically proposed more offers than both agents, with the KB agent usually generating twice as many offers as the QO agent. However, even the KB agent on average proposed far fewer offers than people (5.91 versus 6.97 in the work domain, and 7 versus 8.57 in the trade domain). This result confirms that the QO agent did in fact on average propose fewer offers than the KB agent, but even the KB agent leaves room for improvement in generating more offers.

Work Domain	QO Agent	KB Agent	Human
Number of Negotiation Sessions	33	33	47
Average Offers per Session	2.47	5.91	6.97
Tobacco Domain	QO Agent	KB Agent	Human
Number of Negotiation Sessions	33	33	54
Average Offers per Session	3.53	7	8.57

Table 6 The average number of offers made by the QO agent, KB agent, and People in the Employer / Employer Negotiation Domain (first two lines) and in the Trade Negotiation Domain (last two lines).

These results lead us to the following three conclusions about creating improved negotiation agents. First, bounded rationality, and particularly the tenets of AAT should be explicitly used by agents to model their interactions. Second, the agent should accurately approximate the aspiration scales of the people with whom they interact – something the QO agent was more successful in doing than the KB agent. Finally, automated negotiation agents need to have a mechanism for generating new possible agreements, again based on the aspiration scales of the people with whom they interact. This was something the KB agent was more successful in doing than the QO agent. We leave the specifics of the implementation of this approach and its testing for future work.

6 Conclusions and Future Work

In this paper we report strong empirical findings that bounded rationality, and specifically AAT should be used to model agents designed to operate in optimization and negotiation domains. We studied hundreds of agents written to operate in two general optimization problems. These problems can be generalized to many real-world domains [1,26]. We also studied two state of the art automated negotiation agents [11,19], and hundreds of their interactions with people within two negotiation domains involving 5 and 6 parameters.

This paper contributes several significant findings for Artificial Intelligence researchers. First, we empirically demonstrate from the optimizing domain that people, or even the agents they write on their behalf, are poor optimizers. Even when we explicitly asked two different groups of over 70 people to write agents to optimally solve a problem, and an optimal solution existed, they instead chose to use approaches that fell well short of optimal behavior. Thus, one must conclude that encapsulating human behavior based on optimal strategies is not effective for certain domains. Similarly, the negotiation agents were intentionally written to avoid optimal equilibrium strategies, as they found this approach did not yield productive agreements with people [9,11]. Second, we find that key elements of Aspiration Adaptation Theory (AAT) do effectively encapsulate both the optimizing and negotiation agents studied. We found this model to be more accurate than other bounded rationality theories [2] or traditional machine learning techniques [21] in describing the behavior of the agents. Finally, this empirical study demonstrates the generality of AAT. AAT was originally formulated for domains where utility cannot be measured and thus did not provide any guarantees about performance, or how close to optimal this behavior would be [23]. Thus, the results in this paper are particularly important, and indicate the applicability of AAT even in problems where optimal solutions exist. Along these lines, we believe that AAT can and should be used to design more effective agents for simulation and game environments. In the previous section we outlined several guidelines for this direction, and we hope to study this issue in the future.

For future work, several additional directions are possible. First, while we found that people's strategies fell short of optimal behavior in the optimization problems we studied, we assume that people will write rational and optimal agents in simpler problems. We hope to study the level of problem complexity that motivates people to abandon optimal solutions for those based on bounded rationality. In this paper, we used the strategy method which assumes that people were able to quantify their own strategies. We also assume that as tasks become progressively more complex people may no longer be able to quantify their own strategies, and the fundamental assumption of the strategy method may fail. We also hope to study at which level of problem complexity this occurs. Second, we hope to study how effective the above general conclusions are in simulating human behavior in these and other domains. Finally, we also would like to study how people interact with agents based on AAT versus those based on optimal or other predefined heuristic strategies. Specifically, we hope to study agent-human interactions in an emergency response domain. We are hopeful that AAT and other theories of bounded rationality can be applied to these and other agent-based domain problems.

References

1. Michal Chalamish, David Sarne, and Sarit Kraus. Programming agents as a means of capturing self-strategy. In *AAMAS '08*, pages 1161–1168, 2008.
2. Gerd Gigerenzer and Daniel G. Goldstein. Reasoning the fast and frugal way: models of bounded rationality. *Psychology Rev.*, 103(4):650–669, 1996.
3. P. Terrence Hopmann. *The Negotiation Process and the Resolution of International Conflicts*. University of South Carolina Press, Columbia, SC, May 1996.
4. Leslie P. Kaelbling, Michael L. Littman, and Anthony R. Cassandra. Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101(1-2):99–134, May 1998.
5. Daniel Kahneman and Amos Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47:263–291, 1979.
6. Patrick Kenny, Arno Hartholt, Jonathan Gratch, William Swartout, David Traum, Stacy Marsella, and Diane Piepol. Building interactive virtual humans for training environments. In *Proceedings of Interservice/Industry Training, Simulation and Education Conference (IITSEC)*, 2007.

7. Gregory E. Kersten and Sunil J. Noronha. www-based negotiation support: design, implementation, and use. *Decision Support Systems*, 25(2):135–154, 1999.
8. Claudia Keser and Roy Gardner. Strategic behavior of experienced subjects in a common pool resource game. *International Journal of Game Theory*, 28(2):241–252, 1999.
9. Sarit Kraus, Penina Hoz-Weiss, Jonathan Wilkenfeld, David R. Andersen, and Amy Pate. Resolving crises through automated bilateral negotiations. *Artificial Intelligence*, 172(1):1–18, 2008.
10. Raz Lin and Sarit Kraus. Automated negotiations: Can automated agents proficiently negotiate with humans? In *Communications of the ACM*, page to appear.
11. Raz Lin, Sarit Kraus, Jonathan Wilkenfeld, and James Barry. Negotiating with bounded rational agents in environments with incomplete information using an automated agent. *Artificial Intelligence*, 172(6-7):823–851, 2008.
12. Raz Lin, Yinon Oshrat, and Sarit Kraus. Investigating the benefits of automated negotiations in enhancing peoples negotiation skills. In *Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 345–352, 2009.
13. Pattie Maes. Artificial life meets entertainment: lifelike autonomous agents. *Commun. ACM*, 38(11):108–114, 1995.
14. Efrat Manisterski, Raz Lin, and Sarit Kraus. Understanding how people design trading agents over time. In *AAMAS '08*, pages 1593–1596, 2008.
15. Yohei Murakami, Kazuhisa Minami, Tomoyuki Kawasoe, and Toru Ishida. Multi-agent simulation for crisis management. In *IEEE Workshop on Knowledge Media Networking (KMN02)*, 2002.
16. Yohei Murakami, Yuki Sugimoto, and Toru Ishida. Modeling human behavior for virtual training systems. In *AAAI*, pages 127–132, 2005.
17. John Von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 1944.
18. Martin J. Osborne and Ariel Rubinstein. *A Course In Game Theory*. MIT Press, Cambridge MA, 1994.
19. Yinon Oshrat, Raz Lin, and Sarit Kraus. Facing the challenge of human-agent negotiations via effective general opponent modeling. In *Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 377–384, 2009.
20. David V. Pynadath and Milind Tambe. The communicative multiagent team decision problem: Analyzing teamwork theories and models. *JAIR*, 16:389–423, 2002.
21. Stuart J. Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach (Second Edition)*. Prentice Hall, 2003.
22. Heinz Sauermann and Reinhard Selten. Anspruchsanpassungstheorie der unternehmung. *Zeitschrift für die Gesamte Staatswissenschaft*, 118:577–597, 1962.
23. Reinhard Selten. Aspiration adaptation theory. *Journal of Mathematical Psychology*, 42:1910–214, 1998.
24. Reinhard Selten, Klaus Abbink, Joachim Buchta, and Abdolkarim Sadrieh. How to play (3 x 3)-games.: A strategy method experiment. *Games and Economic Behavior*, 45(1):19–37, October 2003.
25. Reinhard Selten, Michael Mitzkewitz, and Gerald R. Uhlich. Duopoly strategies programmed by experienced players. *Econometrica*, 65(3):517–556, 1997.
26. Reinhard Selten, Sabine Pittnauer, and Martin Hohnisch. Experimental results on the process of goal formation and aspiration adaptation. In *Technical Report – Bonn Laboratory for Experimental Economics*, 2008.
27. Herbert A. Simon. *Models of Man*. John Wiley & Sons, New York, 1957.
28. Moshe Tennenholtz. On stable social laws and qualitative equilibrium for risk-averse agents. In *Proceedings of the 5th International Conference on Principles of Knowledge Representation and Reasoning (KR-96)*, pages 553–561, 1996.
29. Ernest M. Thiessen, Daniel P. Loucks, and Jerry R. Stedinger. Computer-assisted negotiations of water resources conflicts. *GDN*, 7(2):109–129, 1998.
30. Sebastian Thrun. Robotic mapping: A survey. In *Exploring Artificial Intelligence in the New Millenium*. Morgan Kaufmann, 2002.
31. Martin L. Weitzman. Optimal search for the best alternative. *Econometrica*, 47(3):641–654, 1979.
32. Jonathan Wilkenfeld, Kathleen Young, David Quinn, and Victor Asal. *Mediating International Crises*. Routledge, London, 2005.

A Instructions given to the participants

A.1 Commodity Search Commodity Directions

You have decided to buy a new TV to watch the world-cup finals tonight at 9PM. It is now 8AM. You plan to take some time off work to check prices and close a deal. You know the exact brand and the model you are interested in, however, unfortunately, none of the stores will give you their real price (after negotiation) over the phone. In order to learn real prices, you will have to physically enter the store and negotiate. You do know, however the distribution of prices which is normal, mean=\$1000 and standard deviation=200, from the consumer report. Traveling to each store (and in-between stores) and negotiating takes 30 minutes. Once you buy the TV, it will be delivered to your house within 1 hour - wanting to make it to the world-cup finals, you must close the deal by 8PM. If at some point you decide that you want to buy the TV in one of the stores you already scanned, then you can simply call that store, make an order and end the search. Principally, you don't mind spending the entire day searching, however the earlier you buy, the sooner you will be able to get back to work. Note that your hourly salary is \$30 and you lose the hourly pay when you spend your time looking for the TV.

An implementation of this environment can be found at: <http://www.jct.ac.il/~rosenfa/costSearch.html>

After you have played the simulation, you are required to write an agent to purchase the commodity according to your own strategy. In other words, the agent's strategy must be implemented in Java and be consistent with the strategy you formulated for yourself in the first stage. Your grade will be dependent on how closely your agent mimics your own logged games.

You are provided a fully functional Java skeleton implementation of the environment. The only element you must complete is your own strategy within the Agent.java file. You will be provided javadoc files which include detailed information about the classes and functions in the skeleton code.

You are only to submit the code from your Agent.java file. Do not make any changes to the other simulator files. If you do, these changes will not be run. In addition to your agent, you must submit a Word document which includes the following details:

1. A plain text description of your agent.
2. A description of any data structures or variables your agent uses.
3. A detailed description of any methods you added.

The Agent.java file implements one function, decide. This function accepts parameters with information about the current store's commodity price, the store's number, the time elapsed so far, and the time remaining. The function returns the agent's decision, either to continue the search, or to stop the search and buy the commodity in one of the stores already visited.

A.2 Monopoly Domain Directions

Consider an environment of a company with a monopoly for a certain product. The owners of the company must set several important variables that will impact how much money the company will make. In this environment, there are no external factors to these decisions. Thus, the outcome of these decisions is not influenced by any external factors such as how other companies perform, what were other people's decisions, or random effects (all decisions are deterministic).

The task goal is to achieve the highest possible amount of profit for this company over a time period of 50 cycles (also called the company's total profit). The amount of total profit you achieve will be the sum of the profits and losses made over the total 50 periods. No interest (positive or negative) will be taken into account for money made or lost throughout the 50 cycles. Assignment Overview In this assignment, your goal is to write an agent that will control the decisions of the company. While there are no external factors that impact the company's profit or loss, there are many internal parameters that do impact this value. Specifically, there are functions that affect the sales (gross profit) and costs of the company. You are able to control two functions in your company: the product's price and money spent on quality. In general, you need to find the best value for each of these factors. Setting a price function too low will lower profits as each buyer pays too little for the product. Setting a price function too high will lower profit as not enough people will be interested in your product. The product's quality also affects how many people will want to buy the product. In general, the higher the product's quality, the more people will want to buy it. However, the money you spend on quality is deducted as a cost. Thus, some balance must be found in setting this value as well. Java File Overview In this assignment, you are given three Java files: Factory.java, CostSearch.java, and Agent.java. You can download these files from www.jct.ac.il/~rosenfa/monopoly.zip. The files Factory.java and CostSearch.java

are necessary for the experiment setup. You cannot change these files. Your assignment is to best complete the function `calcPrice` and `calcQuality` in the file `Agent.java` such that you maximize your company's profit over the 50 cycles the company exists. Java File Detail File: `Factory.java` This file controls the parameters relating to values of the company's profit and losses. The optimal price in the experiment is based on a normal distribution centered near a value, `price_opt`. The value of this variable in the file you receive is only a sample value. The value for this variable in the trials to evaluate your agent will be set to some value(s) between 25 and 75. The optimal value for the quality function is based on historical quality values and a set of variable weights, `quality_gain1`, `quality_gain2`, and `quality_gain3`. In evaluating your agent, these variables will be kept constant from the values you see in this file. However, the variable, `quality_cost`, which also impacts the cost for quality, and thus also impacts the optimal value you should select for quality, can and will be changed in the evaluation trials (some value between 40 and 60). Due to the quality function used, there are delay affects in the values you set for quality. Note that all of these variables are private and cannot be accessed from the file `Agent.java`.

The values you choose for the agent's price and quality are sent to the functions `setPrice` and `setQuality`. The resulting values from these functions are multiplied to create the company's gross profit. This value is then deducted by the quality costs to create the company's net profit. Note that this file maintains an array (vector array) of the previous values for the price and quality (which can be accessed through the public `getPrice` and `getQuality` functions). You can access the historical profit information for a given cycle through the `getProfit` function. Finally, you can obtain the total profit (total wealth) of a given cycle through the `getWealth(int)` function, or the final total profit through the `getWealth()` function.

File: `CostSearch.java` This file controls the instantiation of the `Agent` object and control of the 50 cycles of an experiment.

File: `Agent.java` This is the file containing the functions `calcQuality` and `calcPrice` that you must supply. Any questions should be sent to Avi Rosenfeld at rosenfa@jct.ac.il.

A.3 Employer / Employee Negotiation Directions

Thank you for participating in the experiment. This document provides explanations on performing the experiment. The experiment itself contains two parts, and each part lasts approximately thirty minutes.

Background:

You are about to take part in a negotiation in a job interview between an employer and a job candidate. The negotiation takes place after the interview is over and both the employer and the job candidate wish to formalize the hiring terms and conditions of the applicant. You will represent the _____ (employer / job candidate).

Both the employer and the candidate must reach an agreement regarding the job description and benefits. The agreement consists of the following issues:

1. **Salary:** The total net salary the applicant will receive. The possible values are: (a) \$7,000, (b) \$12,000, or (c) \$20,000.
2. **Job description:** The job description and responsibilities given to the job applicant. The job description has an effect on the advancement of the candidate in the work place and his/her prestige. The possible values are: (a) QA, (b) programmer, (c) team manager, or (d) project manager.
3. **Social benefits:** The social benefits are in addition to the salary and thus impose an extra expense to the employer, yet provides better terms and conditions for the candidate. The social benefits are divided into two categories: company car and the percentage of the salary allocated by the employer to a pension fund: The possible values for company car are: (a) leased company car, (b) no car, or (c) no agreement. The possible values for % of allocation for pension fund: (a) 0%, (b) 10%, (c) 20%, or (d) no agreement.
4. **Promotion possibilities:** This issue is a commitment by the employer regarding the fast track for promotion for the job candidate. The possible values are: (a) fast promotion track (2 years), (b) slow promotion track (4 years), or (c) no agreement.
5. **Working hours:** This issue is the number of working hours required by the employee per day (not including over-time). This is an integral part of the contract. The possible values are: (a) 8 hours, (b) 9 hours, or (c) 10 hours.

The total time remaining for the negotiation on the contract is 30 minutes, consisting of fifteen 2-minute rounds. If both sides do not reach an agreement by the end of the allocated time, the job interview is over without hiring the candidate. Every side can also opt-out of the negotiation if he feels that the prospects of reaching an agreement with the opponent side are getting slim and it is impossible to negotiate any longer. Opting-out by the job candidate entails the postponement of the project for which the candidate was interviewing, with the possible prospect of its cancelation and a considerable amount of expenses.

Opting-out by the job candidate will make it very hard for him/her to find another job, as the employer will spread his/her negative impression of the candidate to other CEOs of large companies.

Time affects the values of the agreement as well. Every side has a time effect that controls how much the value of an agreement increases (if the time effect is positive) or decreases (if the time effect is negative) as compared to its initial value as the negotiations continue. As time passes the impression of the candidate in the eyes of the employer are less good, and thus the candidate loses, while the employer gains utility, as time passes.

Your goal in the negotiation is to reach an agreement in which your score is the maximum you can receive, as you can deduce from your score table. Please note that the exact score table of your opponent is unknown, though you do know that your opponent can be characterized by one of the following types:

For the job candidate (see score tables):

1. Short-term orientation: the candidate has a family to support and needs the job now. The candidate put much weight on less working hours and some weight on a high salary. The candidate is indifferent to the exact job description, the social benefits and the promotion track.
2. Long-term orientation: the candidate is currently working at another job, he/she has experience in the field and believes he/she has enough experience to improve his/her status in the market place. The candidate puts more weight on receiving a high salary, on a good job description and on a fast promotion track than on the other issues in the negotiation.
3. Compromise orientation: the candidate is willing to compromise now in order to get to a good job in a good company, believing that this can be improved when excelling in the job he/she receives. The candidate wants a commitment for a fast track promotion, while social benefits are not that important.

For the employer (see score tables):

1. Short-term orientation: the employer needs an employee now to help bring the project to a successful ending, without postponing the project's deadline. The main issue for the employer is that the worker will be committed to working many hours and provide a descent salary, while the social benefits are not that important.
2. Long-term orientation: the employer seeks to hire an adequate employee who can be trained as the next generation in the company management. The employer places great importance on the job description and responsibilities of the candidate while he/she is indifferent to the working hours and social benefits.
3. Compromise orientation: the employer is willing to compromise now regarding the financial benefits of the employer, under the assumption that the employer will either prove himself/herself in the future and will even get a raise, or will be mediocre and will not get a raise in the next round of salary raises. However, the employer will not compromise on the working demands from the candidate. Thus, the employer places high importance on the working hours, while he/she is indifferent to the salary and social benefits.

A.4 Political Dispute Negotiation Directions

Thank you for participating in the experiment. This document provides explanations on performing the experiment. The experiment itself contains two parts, and each part lasts approximately thirty minutes.

Background:

You are about to take part in negotiations in a world summit organized by the World Health Organization. The main issue of the negotiations is the Framework Convention on Tobacco Control (FCTC), the world's first public health treaty. The principal goal of the convention is "to protect present and future generations from the devastating health, social, environmental and economic consequences of tobacco consumption and exposure to tobacco smoke." The participating sides in the negotiation are the leaders of England and Zimbabwe. You will represent the leader of _____.

Both leaders must come to an agreement on the following issues:

1. The total amount to be deposited in the Global Tobacco Fund: This issue has an impact on the budget of England and on the effectiveness of economic benefits for Zimbabwe. The possible values are: (a) \$10 billion, (b) \$50 billion, (c) \$100 billion, or (d) no agreement.
2. Impact on other aid programs for Zimbabwe: England may want to pay for the fund by reducing other aid programs. This could affect the net cost of England and the overall benefit for Zimbabwe. Reducing the support in other aid programs by England could have consequences on the socio-economic conditions in Zimbabwe. The possible values are: (a) no reduction, (b) reduction equal to half of the Global Tobacco Fund, (c) reduction equal to the size of the Global Tobacco Fund, or (d) no agreement.

3. Trade policy: Both sides can use trade policy to extract concessions or provide incentives to the other party. There are both benefits and costs involved in these policies: tariffs (taxes on imports) may increase revenue in the short run but lead to retaliation by affected countries. Increasing imports can cause problems for domestic industries. However, it can also lead to lower consumer costs and improved welfare. The possible values are: Zimbabwe's side: (a) Zimbabwe will reduce tariffs on imports, (b) Zimbabwe will increase tariffs on imports, or (c) no agreement. England's side: (a) England will reduce imports, (b) England will increase imports, or (c) no agreement.
4. Creation of a forum to explore comparable agreements for other long-term health issues: If the fund is established, the developing countries will be highly motivated to apply the same approach to other global health agreements. These would be of great benefit to countries suffering from health problems like AIDS and malaria, but would be very costly for England. The possible values are: (a) creation of a fund, (b) creation of a committee to discuss the creation of a fund, (c) creation of a committee to develop an agenda for future discussions, or (d) no agreement.

The total time for the summit is 14 weeks. If both sides do not reach an agreement by the end of the summit, the summit ends and both sides return to the status-quo. Each side can also opt-out of the negotiation if he feels that the prospects of reaching an agreement with the opponent side are getting slim and it is impossible to negotiate any longer. Opting out by England entails imposing sanctions on Zimbabwe by England (including the boycott of tobacco from Zimbabwe), while opting-out by Zimbabwe entails embargoing all British imports to the country. Time affects the values of the agreement as well. Every side has a time effect that controls how much the value of an agreement increases (if the time effect is positive) or decreases (if the time effect is negative) as compared to its initial value as the negotiations continue.

Your goal in the negotiation is to reach an agreement in which your score is the maximum you can achieve, as you can deduce from your score table. Please note that the exact score table of your opponent is unknown, though you do know that your opponent can be characterized as one of the following types: (a) short-term orientation (see score table); (b) long-term orientation (see score table); (c) compromise orientation (see score table).