Using Aspiration Adaptation Theory to Improve Learning*

Avi Rosenfeld[1] and Sarit Kraus[2]
[1]Department of Industrial Engineering
Jerusalem College of Technology, Jerusalem, Israel 91160
[2]Department of Computer Science
Bar-Ilan University, Ramat-Gan, Israel 92500
rosenfa@jct.ac.il, sarit@cs.biu.ac.il

ABSTRACT

Creating agents that properly simulate and interact with people is critical for many applications. Towards creating these agents, models are needed that quickly and accurately predict how people behave in a variety of domains and problems. This paper explores how one bounded rationality theory, Aspiration Adaptation Theory (AAT), can be used to aid in this task. We extensively studied two types of problems – a relatively simple optimization problem and two complex negotiation problems. We compared the predictive capabilities of traditional learning methods with those where we added key elements of AAT and other optimal and bounded rationality models. Within the extensive empirical studies we conducted, we found that machine learning models combined with AAT were most effective in quickly and accurately predicting people's behavior.

Categories and Subject Descriptors

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

General Terms

Experimentation

Keywords

bounded rationality, cognitive models, agent learning

1. INTRODUCTION

The challenge of creating agents that effectively simulate and interact with people is of utmost importance in many systems [7, 8, 9, 16]. These agents form the backbone of many mixed human-agent systems such as entertainment domains [7], Interactive Tutoring Systems [9], and mixed human-agent trading environments [8]. In order to effectively interact with people, the agents have to understand and predict people's behavior. To date, these agents have often been created based on the perspectives of unbounded rationality including expected utility, game theory, Bayesian models, or Markov Decision Processes (MDP) [10, 15].

While these models are mathematically elegant and have proven effective in some situations [10, 15], several fundamental obstacles exist in applying them to many real-world applications. 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 [3] 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. Thus, even assuming that 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 [12]. Thus, even in the best of circumstances, expecting people to behave optimally based on full rationality is unrealistic.

We posit that models based on Bounded Rationality hold the most promise to best predict people's behavior. This research direction, initiated by Simon [17], assumes that people – except in the simplest of 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) [16] as a boundedly rational model of decision making.

This paper's major contribution is support for the claim that learning models that aim to predict people's behavior should be based on bounded rationality models, and specifically AAT. To empirically support this claim, we studied two types of problems – one relatively simple optimizing problem and two complex negotiation problems. Within the optimization problem, we found that traditional machine learning methods such as decision trees were successful in discovering the strategies most people used. We note that these strategies were consistent with AAT and were not representative of other bounded theories or the problem's optimal policy. Furthermore, we found that an AAT based predic-

*This research is based on work supported in part by the U.S. Army Research Laboratory and the U.S. Army Research Office grant # W911NF-08-1-0144, NSF grant # 0705587 and the Israel Ministry of Science and Technology grant # 3-6797. Sarit Kraus is also affiliated with UMIACS.

Cite as: Title, Author(s), *Proc. of 10th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2011)*, Tumer, Yolum, Sonenberg and Stone (eds.), May, 2–6, 2011, Taipei, Taiwan, pp. XXX-XXX.

tion policy was extremely accurate, even when trained with very limited data. Second, in the negotiation domains we studied, traditional learning algorithms such as those in the Weka learning package [21] were unable to effectively predict how people would behave due to the inherent problem complexity. Nonetheless, by adding statistical information about people's bounded AAT strategies, we were able to significantly improve models' prediction accuracy beyond the base algorithm.

2. RELATED WORK

Many multi-agent researchers have found that classical models based on rigid models of expected utility often are not effective in domains they studied. [1, 4, 5, 6, 8, 14]. Consequently, a key research question is what should be used in place of these models. For example, designers of automated negotiation agents found that when the automated agents follow their equilibrium strategies, the human negotiators who negotiate with them become frustrated as the equilibrium strategy requires the automated agent to repeatedly propose the same offer. Thus, negotiation sessions often ended with no agreement [4]. Similarly, Lin et al. [6], created an equilibrium agent for the negotiation domains studied in this paper. They also found that in many situations the equilibrium agent failed to reach any agreement when playing with people [6]. Thus, both the KB and QO agents they created also intentionally steered away from the equilibrium strategies [6]. In fact, a survey article of automated negotiators [5] found that none of the reviewed agents implemented equilibrium strategies. They attributed this to the assumption that people avoid equilibrium strategies in complex environments due to their bounded nature.

However, once classical rationality models have been rejected, the question then becomes what model to use instead. Towards this goal, previous works typically model people's preferences based on historical information. For example, Gal and Pfeffer use a statistical approach to learn which bid a person will select from a known number of possibilities [1]. Pardoe and Stone used machine learning techniques to create a trading agent to predict the likelihood that a person would accept a given bid price [11]. However, a key methodology difference exists between these and works similar to ours. While previous works typically used existing machine learning algorithms, our goal is to find a general model of bounded rationality and use it to better learn people's actions. This important difference impacts two key research issues. First, by using only generalized bounded theories, we can prove or disprove the applicability of psychological or economic bounded models in new problem domains. Possibly more importantly, and as we demonstrate in this paper, once a generalized theory is found to be accurate, it can be applied to new problems in order to further improve existing learning algorithms' accuracy above traditionally used machine learning methods.

Previously, we studied the applicability of bounded rationality models, and even AAT [14]. However, that study focused on validating if AAT was relevant, and not how it could be used for improving learning accuracy. Additionally, that work used human judges to decide if AAT was being used, and thus leaves open concerns regarding possible human bias. This paper is the first that uses classic decision tree models [13] to decide what decision model, bounded or not, is used and how to use these theories for improved learning.

3. ASPIRATION ADAPTATION THEORY

Aspiration Adaptation Theory (AAT) was proposed by Selten as a general economic model for how people make certain economic decisions without any need for expected utility functions [16]. AAT was originally formulated to model how people make decisions where utility functions cannot be constructed. For 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? While in theory utility based models could be used, many of us do not create rigid formulas involving numerical values to weigh trade-offs between each of these search parameters.

AAT provides an alternative to utility theory for how decisions can be made in this and other problems. First, m goal variables are sorted in order of priority, or their *urgency*. Accordingly, the order of $G_1, \ldots, G_m$ refers to goals' urgency, or the priority by which a solution for the goal variables is attempted. Each of the goal variables has a desired value, or its *aspiration level*, that the agent sets for the current period. This desired valued 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 starts 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.

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, within traditional learning or search, optimal values for all variables are sought for simultaneously [15]. In contrast, within AAT only one goal is attempted to be satisfied 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.

It is important to note that two key differences exist between classic AAT, and how we apply AAT within this study. First, AAT assumes that the m goal variables used to solve $\mathcal{G}$ are incomparable as no utility function is possible to connect goal variables. For example, in buying a house the goal variables for location, price and size are likely to be incomparable with no evident utility function to compare them. In this paper, we consider simpler 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 their bounded nature. This represents a significant generalization to AAT's theory and its relevance even within 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 optimization and negotiation problems were utility can be calculated, 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. This difference again represents a significant generalization of AAT.

We recognize that AAT is not the only possible model that can predict human decisions. In both of the problems we studied, optimal search methods or negotiation strategies could have been used. Additionally, other 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, 3]. Following the psychological approaches, simple predefined or greedy heuristics could be used. In the optimizing domains, predefined values could have been used instead of attempting to solve the problems. Within 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".

4. EXPERIMENT SETUP

We studied how people solved two types of problems – a relatively simple optimization problem and more complex negotiation problems. We used the optimization problem as a baseline as we believe it would be easier to understand the decision making models used by people in this problem. By focusing on more complex problems as well, we show the applicability of our findings to real-world applications.

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 [20]. The use of simulation and role-playing is common for training people in negotiations (e.g., the Interactive Computer-Assisted Negotiation Support system (ICANS)) [18]. 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.

4.1 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. Assume 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 [19] 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 $\mathcal{T} = \{1, 2, ..., k\}$. 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: exists $\bar{c}$ such that if $c_t \leq \bar{c}$ the agent should stop the search [19].

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 was 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) [19].

In our implementation, the prices are distributed normally with a mean μ and a standard deviation σ. We denote by x 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 probability[1] for p is

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

Given these definitions we must generally solve:

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

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

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

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

4.2 Negotiation problems

We also studied two previously defined negotiation domains [6] and focused on how people came to agreements with other people in these problems. The goal of these problems was to negotiate as high a utility as possible given a known weight of all issues. To reach an agreement, people sent offers through a web interface which facilitated their choosing the different values that constitute an offer. This offer was then sent to the other person in plain English. A time effect existed that assigned a time cost which influences

[1]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 consider it in our analysis.

the utility of each player as time passed (there can be different time costs for each player). The time effect was either negative or positive. If no agreement is reached by the end of the final turn then a status quo agreement was implemented resulting in a status quo value for each player. Each player could also quit the negotiation session at any given time if he/she decided that the negotiation session is not proceeding in a favorable way. This resulted in the implementation of an opt out outcome. During each phase of the negotiation session, the instructions and parameters subject to negotiation were accessible to the players. The players were 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 was also easily accessible. When receiving an offer the player can choose whether to accept or reject it, or make a counter-offer.

4.2.1 Employer / Employee Negotiation Domain

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: The **Salary** 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. The **Job Description** 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. In addition to the base salary, other job benefits may also be negotiated. The **Car Benefits** issue revolves around the possibility that the company will provide a company car for use by the employee with possible values being (a) a leased company car or (b) no leased car. Pension benefits must also be negotiated and set as percentage of the work's salary. The possible value for the **Pension benefits** are (a) 0%, (b) 10%, or (c) 20%. The **Promotion possibilities** 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), or (b) slow promotion track (4 years). The **Working hours** 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. In total, this scenario allows for a total of 1,296 possible agreements $(3 \times 4 \times 12 \times 3 \times 3 = 1296)$[2].

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.

[2]Note that it is possible for there to be no agreement for many of these parameters. In these case, we considered the negotiation to have failed.

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.

4.2.2 Political Dispute Negotiation Domain

The second negotiation is based on a scenario where England and Zimbabwe attempt to reach an agreement evolving from the World Health Organization's 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: The **Creation of a Global Tobacco Fund** 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, or (c) \$100 billion. The **Impact on other aid programs** 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, or (c) reduction equal to the size of the Global Tobacco Fund. Both Zimbabwe and England must negotiate **Trade Issues**. Countries 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 both Zimbabwe and England possible values for this are: (a) reducing tariffs on imports or (b) increasing tariffs on imports. The **Forum to Study Long-Term Health Issues** 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, or (c) creation of a committee to develop an agenda for future discussions. Consequently, a total of 576 possible agreements exist ($4 \times 4 \times 3 \times 3 \times 4 = 576$)[3].

5. EXPERIMENTAL RESULTS

In this section we detail the shortcomings of using optimal models to predict people's behavior, and how AAT, and not other bounded methods, should be used instead. In general, we found that in the optimization and negotiation problems we studied, key elements of AAT were present in decisions people made. When sufficient data was present, as we found was the case in the relatively simple optimization problem, clear strategies consistent with AAT were learned by standard machine learning algorithms such as C4.5 [13]. We found that estimates of people's aspiration scales in this problem were useful for formulating a very accurate prediction model even without an extended learning period and with only sparse data. Within the more complicated negotiation problems, we found that adding statistical information about people's typically aspiration scales was critical for improving the prediction models as using both the equilibrium strategies and traditional learning methods yielded an extremely poor predictor of how people would act.

5.1 AAT to Predict Optimization Decisions

Our first goal was to use machine learning techniques to determine if optimal polices, fast and frugal heuristics [2] or AAT best predict people's optimizing decisions. Recall from Section 3.1 that an optimal policy exists based on price alone – buy if the price in the current store is less than 789. Thus classical expected utility theory would predict that people would similarly buy the commodity at this price. Assuming people used fast and frugal search heuristics, we would expect them to formulate simple strategies involving only one variable (e.g. search until price $< X$, or visit Y stores and buy in the cheapest store). However, using an AAT based model for prediction would assume some type of combination strategy exists where one variable is first searched for, but then retreated from assuming that value could not be satisfied. For example, a person might initially search for a price less than 650, but will settle on even a higher price (e.g. the lowest found so far) after unsuccessfully finding this price after 5 stores.

We did in fact find strong support that AAT best predicted when people would buy the commodity. To study this point, we studied the log files taken from 41 people and how they chose to buy the commodity. Each person was presented with a simulated implementation of the problem described in Section 3.1[4]. Every interaction with the simulation randomized the values for the commodity as described in Section 4.1, and all people were told to interact with this simulation until they formulated a clear policy for how they would decide to buy the commodity. After this point, we then logged a minimum of 20 additional interactions (average 25.56) where each interaction ended in a decision to buy the commodity in a certain store. Our goal was to predict where each person would end his search process.

To obtain a prediction model without bias, we first separated all logged interactions for each of the 41 people, entered them as input for the Weka machine learning package [21], and applied the C4.5 decision tree classification algorithm as implemented by Weka to decide if each person's behavior was consistent with AAT. We chose this learning algorithm, as opposed to others such as Bayes or Neural Nets, as the output from the C4.5 algorithm would not just predict the person's behavior, but also provide the rules by which the classifier operates. We could then judge if these rules were consistent with the optimal rule (e.g. purchase if price < 789) or if the classifier represented a fast and frugal rule with only one rule (e.g. buy after 4 stores). Weka found that 30 of the 41 strategies had decision rules based on price and store combinations (e.g. buy immediately if the price is less than 700, otherwise visit 5 stores and buy in the cheapest one) which are clearly non-optimal and not frugal. Instead, this decision process can be viewed as a classic example of the urgency and retreat process within AAT. In these strategies, a certain price threshold is desired (highest urgency) but retreated from if believed to be unattainable. Of the remaining 11 rules, 10 rules were found to be based on the price variable alone, with one based on the number of stores. While none of these 11 rules were optimal, they may be viewed as fast and frugal heuristics. Thus, the C4.5 learning algorithm found that the majority of strategies (30 of 41 or 73%) were consistent with urgency and retreat concepts of AAT, while the minority (27%) of the strategies were an alternate bounded rationality model – namely simpler fast and frugal heuristics. None of the strategies were found to be optimal. This result provides strong support to the claim that AAT, and not optimal or fast and frugal heuristics best predict people's behavior. Also note that these results are consistent with our previous work [14] where human judges were used instead of machine learning techniques.

Next, we wished to create a general prediction model and check how AAT might be used to improve such a model. To create and test such a model, we combined all 41 people's logged interactions to create a total of nearly 5000 instances where people either decided to buy the commodity or to continue their search. Using this data, we constructed two baseline models, found in Figure 1. One baseline is a *Naive* model that classifies all decisions based on the majority class, here assuming people will always continue the search. As people didn't typically buy the commodity right away, the majority decision is to continue the search and thus 78.56% of all decisions are of this type (see Column 1). A second baseline class is a *Learning* decision tree model constructed which was trained using Weka's C4.5 classifier on the combined data and tested with cross-validation (Column 2). It is interesting to note that this decision tree was also consistent with AAT, but this result is not surprising as most individual logs were consistent with AAT as well.

We then compared these baselines to models which contained adding information about people's AAT preferences. Column 3 of Figure 1 is based on a learned C4.5 model, but added information about people's average AAT preferences (e.g. buy immediately if the price is less than 750, otherwise settle on the best price after visiting a total of 3 stores). While this model is slightly better than the learned baseline, this model was not significantly more accurate. We hypothesized this is due to learned baseline being already based on AAT. Thus, by adding additional information we did not significantly aid the C4.5 to improve its accuracy.

[3]We again consider the negotiation as having failed if no agreement is reached on all issues.

[4]A web interface for this problem can be seen at: http://www.jct.ac.il/~rosenfa/costSearch.html

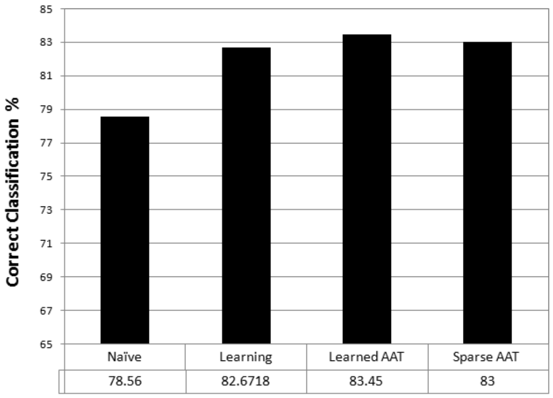


Figure 1: Comparing the Prediction Accuracy between AAT and non-AAT Based Models

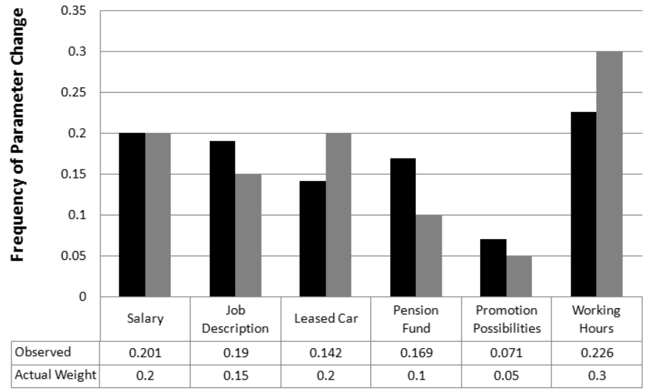


Figure 2: Frequency of Parameter Change in Employer / Employer Negotiation Domain

We hypothesized that when AAT preferences are clear, such as seems the case in this problem, even sparse data could be used to form an accurate model. In this problem, the most urgent goal variable is the commodity price in the current store. By sampling only a limited number of instances where people stopped the search based on this value, we believe it is possible to approximate the accuracy of the learned model from nearly 5000 logged instances. To support this claim, we formed a learning policy based on the average price used in 50 decisions to buy the commodity and simply averaged the threshold for this parameter (average search stopped at price = 767). The accuracy of this policy is presented in column 4 of Figure 1. Observe that this model is nearly equally accurate to the learned policies. In contrast, creating traditional models with such small amounts of data were not successful and yielded the naive model (Column 1). We also observed that even extremely small samples of only 10 decisions formed similar models. We took 5 such samples and noted that the deviation between these samplings was not great and the lowest prediction accuracy was 81.92%. Thus, we conclude that in relatively basic domains, an accurate learning model can be made even with only limited AAT data based on the most important search parameter(s).

5.2 AAT to Predict Negotiation Decisions

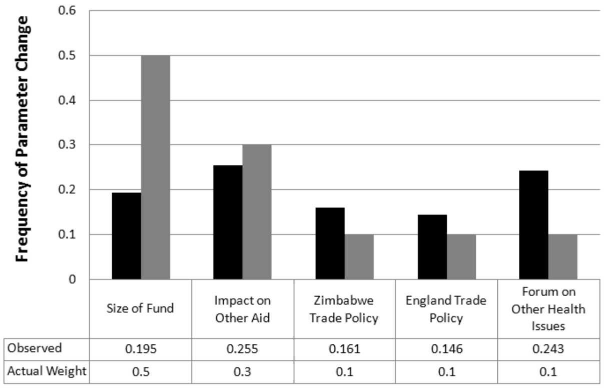
In order to study more complex, real-world problems, we also studied if AAT could be used to better predict people's negotiation activities in the two problems described in Section 3.2. Recall that in these problems people must negotiate either 5 or 6 parameters. In this section, we study two key issues: 1) Is AAT applicable to negotiation problems? 2) Assuming AAT is applicable, will it facilitate a better prediction model for people's behavior?

According to AAT, one would expect people to rank the importance of each of the negotiation parameters according to his or her aspiration scale. Assuming people often have the same aspiration scales, we would also see an order where issues are addressed, e.g. certain parameters are typically negotiated first, second, etc. For example, in the employer / employee domain, we might find that negotiations first focus on the salary amount parameter and only then move on to other parameters such as pension or transportation benefits. Our premise is that by understanding these scales, one can add this information into traditional models such as C4.5 to more accurately predict what bids people will offer.

We did in fact find that aspiration scales existed whereby certain negotiation parameters were stressed more frequently than others. To study this point, we analyzed how frequently the given parameters were changed, on average, over the course of all collected negotiation interactions. As a baseline, we also considered the actual weights these issues were given [6]. This allowed us to compare if people stressed negotiation issues differently from this baseline value.

Figure 2 presents how frequently each of the 6 parameters in the employer / employee domain were changed. These results were taken from 47 negotiation sessions between people. The first column in this Figure shows the frequency people changed each of these issues (either a raised or lower value). The second value represents the actual weights these issues had. Note that clear aspiration scales existed, and these scales were different from the actual issues weights. Certain parameters, say promotion possibilities, clearly had a lower urgency as these issues were typically not discussed. In comparison, other parameters, such as working hours and salary, were discussed frequently. Furthermore, we observed that even within issues such as working hours and salary that seemed to be equally discussed, the point where they entered in negotiation was often not the same. Often the salary point was discussed first, and only then the number of hours. For example, within the first two negotiation interactions, the salary issues was discussed in 45% of all session, while the number of hours issues was only discussed in 24% of the sessions. Again, this would represent that the salary issue had a higher urgency at the start of negotiations.

Figure 3 presents how frequently each of the 5 parameters in the tobacco trade domain were changed. These results were taken from 56 negotiation sessions between people. Again, certain issues such as the Impact on other Aid and Forum on Other Health Issues were clearly discussed more frequently than issues such as England and Zimbabwe's Trade Policy. We again noted that certain issues were typically negotiated at different points of sessions. Also note that the aspiration scales here differ greatly from the actual weights of the issues. The actual weight for the Size of Fund parameter was equal to all of issues combined (0.5 of the total weight). Yet, people typically focused on other issues more, such as the Impact on Other Aid and Forum on Other Health Issues parameters. Thus, deriving aspirations from



	Size of Fund	Impact on Other Aid	Zimbabwe Trade Policy	England Trade Policy	Forum on Other Health Issues
Observed	0.195	0.255	0.161	0.146	0.243
Actual Weight	0.5	0.3	0.1	0.1	0.1

Figure 3: Frequency of Parameter Change in Trade Negotiation Domain

the actual utility weights is not possible.

We then proceeded to study if adding AAT information was helpful in predicting how people will negotiate. In the problems we considered, the parameters to be negotiated could have between 2 and 4 discrete values. Please refer back to Section 3.2 for a description of these parameters and their possible values. The goal of the models was to accurately predict what changes, if any, would be made by a person in the next offer.

In order to study this point we considered several models for both negotiation problems (see Tables 1 and 2). The goal of all of these models was to predict the next value for each parameter. In the learning models, each parameter was training and tested separately through cross-validation, but did have access to the previous values for all parameters. First, we considered the **Majority Rule** model. Given the full log file, this rule assumes that a person would offer the most popular value for any given parameter. For example, in the employer / employee domain, the most popular **title** was "Programmer". Second, we implemented two models based on the **equilibrium strategy**. This strategy is based on previous work in these problems [6]. However, as the equilibrium strategy will change based on which person is allowed to offer the last bid, we checked both what equilibrium strategies would predict for all parameters. Next, we created a baseline strategy that uses the **C4.5** algorithm to predict the next offer for each parameter. This model used historical information about the previous offer and the current negotiation iteration. Next, we created a **C4.5 with AAT statistical information** prediction model. As we previously demonstrated, each parameter had different urgencies. Thus, we attempted to create a more accurate model by adding information about which parameters were typically raised or lower for any given iteration. Specifically, we added a field with a binary flag value to differentiate between the iterations for which people typically changed a given parameters' value with a frequency of ≥ 0.5, and those which were typically not changed and added information would likely not help. This was done to avoid overfitting the AAT statistics for any training / testing pair, and to thus keep the generality of the results. Finally, we created a **C4.5 + Complete Behavior Knowledge** model. This final baseline had knowledge about what the previous offer was, and also added perfect knowledge if the person would revise upwards, downwards, or leave unchanged their previous offer. In cases where only two options exist, one would expect this baseline to guarantee 100% accuracy. However, when more than 3 values exist for a given parameter, even this model cannot guarantee 100% accuracy. For example, if a previous salary offer was $7,000 per month and we know the next offer will be higher, we still do not know if it will be raised to $12,000 or $20,000. Nonetheless, the goal of this model was to provide an upper bound for how much AAT based information could theoretically help.

Tables 1 and 2 demonstrate the effectiveness of adding AAT information to boost prediction accuracy. The first row of these tables show the parameter to be negotiated and the number of possible values. The second row presents the majority rule baseline. The third and fourth rows present how effective the equilibrium policies were in predicting what people actually offered. Note that both of these policies in both problems fall well below the naive majority baseline. This again demonstrates the ineffectiveness of using equilibrium theoretical policies to predict how people actually behave. The fifth row presents the accuracy of the learned C4.5 model. This model represents the effectiveness of this traditional learning method in predicting each of the parameters. We then added AAT information, and reran the same C4.5 algorithm, the results of which are in the sixth row. Note that in both domains the improvement gained from the AAT information is significant. However, in the Tobacco Trade Domain (Table 2), the prediction improvement is much larger from the base C4.5 algorithm (over a 10% accuracy boost for many parameters), yet falls short of the accuracy in the Employer / Employee Work Domain (Table 1). Also, in both domains, only one parameter did not gain from the added aspiration information. For both of these parameters, few instances existed where people had clear general aspiration changes, preventing any accuracy boost from this approach. Finally, the last line in both domains presents the accuracy of the C4.5 algorithm with complete behavior knowledge, or perfect information about whether a person will retreat from (decrease) a given parameter value, or upwardly revise its aspiration (increase). Note that as expected even complete AAT information could not yield 100% prediction accuracy for parameters with more than 2 values.

6. CONCLUSIONS AND FUTURE WORK

This paper makes several significant contributions towards creating more effective agents to interact with people in optimization and negotiation problems. First, we found that "traditional" rationality models were often poor indication how people will act. Consistent with previous research [1, 11], we found that a better alternative is to use traditional learning techniques to predict how people will behave. However, in contrast to previous works, we used decision trees [13] to formulate exactly which policy was used instead. This classifier found that bounded rationality theories, and specifically AAT, were used. Second, this paper represents a unique approach where this general theory was then reapplied to improve learning models. Within the complex negotiation domain, this approach significantly improved prediction accuracy, often by over 10%. Within the simpler optimization problem, this approach was useful in producing accurate learning models even when extremely limited learning data was available.

For future work, several directions are possible. First,

	Salary-3	Title-4	Car-2	Pension-3	Promotion-2	Hours-3	Average
Majority Rule	60.1852	67.5926	57.4074	70.3704	62.963	62.963	63.5803
Equilibrium Strategy 1	44.4444	67.5926	69.4444	66.6667	41.6667	67.5926	59.568
Equilibrium Strategy 2	25.9259	17.5926	69.4444	19.4444	43.5185	61.1111	39.5062
C4.5 Without AAT	61.111	68.5185	68.5185	67.5926	83.3333	69.4444	69.7531
C4.5 with AAT stats	**62.963**	**68.5185**	**75.9259**	**71.2963**	**91.6667**	**76.8519**	**74.53705**
C4.5 + Complete Knowledge	95.3704	89.814	100	96.2963	100	96.2963	96.2962

Table 1: Comparing the Prediction Accuracy between AAT and non-AAT Based Models in the Employer / Employer Negotiation Domain

	Fund Size-3	Aid-3	Zim. Tariff-2	Eng. Tariff-2	Forum-3	Average
Majority Rule	57.3333	45.3333	58.2222	62.6667	44	53.5111
Equilibrium Strategy 1	35.088	42.3849	41.962	61.9443	46.5569	45.58722
Equilibrium Strategy 2	35.088	42.3849	51.337	44.7568	41.9177	43.09688
C4.5 Without AAT	61.8257	46.888	57.7778	64	49.3776	55.97382
C4.5 with AAT stats	**71.1111**	**56.4444**	**67.1111**	**64**	**55.5556**	**62.84444**
C4.5 + Complete Knowledge	95.5556	93.7778	100	100	87.1111	95.2889

Table 2: Comparing the Prediction Accuracy between AAT and non-AAT Based Models in the Tobacco Trade Negotiation Domain

once we have demonstrated that knowing people's aspirations improves learning, one may wish to study how these values can be quickly and accurately identified to further aid in the learning process. Second, this paper focuses on how to best learn how people interact with each other. One important application of this work is to create automated agents that use this paper's lessons to better interact with people. State of the art automated negotiation agents [5, 6] currently do not use this information. Third, we were successful in demonstrating that AAT is more useful than traditional rationality or fast and frugal bounded heuristics to predict how people in the problems we studied. However, an open question is what other, likely bounded, general theories will be helpful in creating learning agents in other real-world environments where AAT might not be relevant.

7. REFERENCES

[1] Y. Gal and A. Pfeffer. Predicting people's bidding behavior in negotiation. In *AAMAS '06*, pages 370–376, 2006.

[2] G. Gigerenzer and D. G. Goldstein. Reasoning the fast and frugal way: models of bounded rationality. *Psychology Rev*, 103(4):650–669, 1996.

[3] D. Kahneman and A. Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47:263–291, 1979.

[4] S. Kraus, P. Hoz-Weiss, J. Wilkenfeld, D. R. Andersen, and A. Pate. Resolving crises through automated bilateral negotiations. *Artificial Intelligence*, 172(1):1–18, 2008.

[5] R. Lin and S. Kraus. Automated negotiations: Can automated agents proficiently negotiate with humans? 53(1):78–88, 2010.

[6] R. Lin, S. Kraus, J. Wilkenfeld, and J. Barry. Negotiating with bounded rational agents in environments with incomplete information using an automated agent. *Artificial Intelligence*, 172(6-7):823–851, 2008.

[7] P. Maes. Artificial life meets entertainment: lifelike autonomous agents. *Commun. ACM*, 38(11):108–114, 1995.

[8] E. Manisterski, R. Lin, and S. Kraus. Understanding how people design trading agents over time. In *AAMAS '08*, pages 1593–1596, 2008.

[9] Y. Murakami, Y. Sugimoto, and T. Ishida. Modeling human behavior for virtual training systems. In *AAAI*, pages 127–132, 2005.

[10] J. V. Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 1944.

[11] D. Pardoe and P. Stone. Bidding for customer orders in tac scm. In *AAMAS 2004 Workshop on Agent Mediated Electronic Commerce VI*, pages 143–157, 2004.

[12] D. V. Pynadath and M. Tambe. The communicative multiagent team decision problem: Analyzing teamwork theories and models. *JAIR*, 16:389–423, 2002.

[13] J. R. Quinlan. *C4.5: Programs for Machine Learning*. Morgan Kaufmann, 1 edition, January 1993.

[14] A. Rosenfeld and S. Kraus. Modeling agents through bounded rationality theories. In *IJCAI-09*, pages 264–271, 2009.

[15] S. J. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall, 2003.

[16] R. Selten. Aspiration adaptation theory. *Journal of Mathematical Psychology*, 42:1910–214, 1998.

[17] H. A. Simon. *Models of Man*. John Wiley & Sons, New York, 1957.

[18] E. M. Thiessen, D. P. Loucks, and J. R. Stedinger. Computer-assisted negotiations of water resources conflicts. *GDN*, 7(2):109–129, 1998.

[19] M. L. Weitzman. Optimal search for the best alternative. *Econometrica*, 47(3):641–654, 1979.

[20] J. Wilkenfeld, K. Young, D. Quinn, and V. Asal. *Mediating International Crises*. Routledge, London, 2005.

[21] I. H. Witten and E. Frank. *Data Mining: Practical Machine Learning Tools and Techniques, Second Edition*. Morgan Kaufmann, June 2005.