Running head: COMBINING PSYCHOLOGICAL MODELS WITH MACHINE LEARNING

Combining Psychological Models with Machine Learning to Better Predict People's Decisions

Avi Rosenfeld[1], Inon Zukerman[3], Amos Azaria[2], 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

[3] Department of Industrial Engineering and Management,

Ariel University Center of Samaria, Ariel, Israel 40700

rosenfa@jct.ac.il, inonzu@ariel.ac.il, azariaa1@mail.biu.ac.il, sarit@cs.biu.ac.il

Abstract

Creating agents that proficiently interact with people is critical for many applications. Towards creating these agents, models are needed that effectively predict people's decisions in a variety of problems. To date, two approaches have been suggested to generally describe people's decision behavior. These models could either be based on theoretical rational behavior, or psychological models such as those based on bounded rationality. A second approach focuses on creating models based exclusively on observations of people's behavior. At the forefront of these type of methods are various machine learning algorithms.

This paper explores how these two approaches can be compared and combined in different types of domains. In relatively simple domains, both psychological models and machine learning yield clear prediction models with nearly identical results. In more complex domains, psychological or machine learning alone cannot accurately predict people's decisions. However, improved models can be created by using machine learning techniques to refine parameters within psychological models. In the most complex domains, the exact action predicted by psychological models is not even clear, and machine learning models are even less accurate. Nonetheless, by creating hybrid methods that incorporate features from psychological models in conjunction with machine learning we can create significantly improved models for predicting people's decisions. To demonstrate these claims, we present a survey of previous and new results, taken from representative domains ranging from a relatively simple optimization problem, a more complex path selection domain, and complex domains of negotiation and coordination without communication.

Combining Psychological Models with Machine Learning to Better Predict People's Decisions

Introduction

The challenge of predicting people's decisions is of utmost importance for many economics, psychologists, and artificial intelligence researchers (Chalamish, Sarne, & Kraus, 2008; Gigerenzer & Goldstein, 1996; Keser & Gardner, 1999; Maes, 1995; Manisterski, Lin, & Kraus, 2008; Murakami, Minami, Kawasoe, & Ishida, 2002; Murakami, Sugimoto, & Ishida, 2005; Selten, 1998; Selten, Abbink, Buchta, & Sadrieh, 2003; Selten, Mitzkewitz, & Uhlich, 1997). Within the field of economics and psychology, validly encapsulating human decision-making is critical for predicting the short and long term effects of a given policy (Neumann & Morgenstern, 1944; Gigerenzer & Goldstein, 1996; Selten, 1998; Kahneman & Tversky, 1979). To computer scientists, accurately predicting people's actions is critical for mixed human-computer systems such as entertainment domains (Maes, 1995), Interactive Tutoring Systems (Murakami et al., 2005), and mixed human-agent trading environments (Manisterski et al., 2008). Within these and similar domains, creating agents that effectively understand and/or simulate people's logic is particularly important.

To date, two approaches have been proposed for predicting people's decisions by social and behavioral scientists. One classic approach, often advocated by economists, has modeled people's behavior based on classic decision theory. This direction, originally proposed by Von Neumann and Morgenstern (Neumann & Morgenstern, 1944) assumes that people's decisions can be modeled mathematically and rationally based on expected utility. Even when people are faced with uncertainty, these models assume people will adhere to strict mathematical formulae based on the probability each event will occur. Game theory follows this approach, and equilibrium strategies, such as the Nash equilibrium (Nash, 1951), apply expected utility to situations where two or more people interact to predict their decisions. These solution concepts have proven effective in some applications (Kaelbling, Littman, & Cassandra, 1998; Neumann & Morgenstern, 1944; Russell & Norvig, 2003). However, research into people's decisions have shown that people do not necessarily always adhere to these rigid models (Gigerenzer & Goldstein, 1996; Selten, 1998; Kahneman & Tversky, 1979).

A second group of approaches, often advocated by psychologists and experimental economists, build cognitive models based on people's subjective perception of a problem. These approaches posit that theoretical outcomes are less important, and models must instead be constructed based on modeling people's observed behavior. Examples of this direction include Kahneman and Tversky Prospect Theory (Kahneman & Tversky, 1979) that models how people deviate from expected utility when faced with risk, and Gigerenzer and Goldstein's fast and frugal heuristics (Gigerenzer & Goldstein, 1996) that assume people use simplistic heuristic to guide their decisions. Models of bounded rationality lie within this group, as they posit that people search 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 (Simon, 1957). We considered one such theory, Selten's Aspiration Adaptation Theory (Selten, 1998), whereby people make decisions by attempting to satisfy only goal variable at a time, or a given "aspiration".

In contrast to both of these cognitive models, computer scientists often model peoples' decisions through machine learning techniques (Russell & Norvig, 2003). These models are based on statistical methods such as Bayes' Rule, Neural Networks, Support Vector Machines (SVM), or Decision Tree algorithms (Mitchell, 1997). These approaches are built exclusively based on observed decisions, instead of generally predicting how people behave. As a result, these models do not make any claims for their general applicability as they were created exclusively based on observations in a specific setting.

The key contribution of this paper is an exploration of how one can combine the decision making approaches proposed by social scientists with classic machine learning approaches. In this paper we present a survey of problems ranging from relatively simple to progressively more complex problems. We refer to the simple problems as those where accurate models are possible from both cognitive and machine learning models. Note that even in these "simple" problems, multiple cognitive models may theoretically be possible allowing us to consider a range of predictions. In a second type of problems, multiple cognitive models are theoretically possible, but due to the complexity of the problem, it is not clear how to apply them. For example, Kahneman and Tversky's Prospect Theory (Kahneman & Tversky, 1979) posits that people are risk adverse and will prefer definite returns. However, this theory does not make specific predictions about parameters within a given problem. While it is clear according to Prospect Theory that most people will prefer 50 Euro over a 50% probability of receiving 100 Euro, would they prefer 49 or 48 Euro over a 50% probability of receiving 100 Euro? In these types of cases, we applied machine learning techniques to discover the parameters within learned models. We found that such combined models were more accurate than those with machine learning alone, yet again, one psychological model was best after this procedure was applied. As problems become progressively more complex, the number of parameters needing to be learned increases, necessitating novel methods for learning these parameters. Due to the complexity of the problem, machine learning methods alone do not perform well. We found that building hybrid models which use as their base machine learning, but add features from psychological models, performed significantly better in these types of problems.

To demonstrate these results, we present a survey of previous and new results. Specifically, we present and discuss how this methodology was applied to several different problems. We present three different psychological models and alternatives, including strictly rational models, in each of the domains that we considered. First, in the next section, we present one bounded rationality model, Aspiration Adaptation Theory (AAT) (Selten, 1998), that describes how people make decisions in the absence of an explicit utility function. We found this model was found to be the best cognitive model in a relatively simple optimization problem, and helped significantly increase machine learning's accuracy in a complex negotiation model. In the following section, we present the Hyperbolic Discount model, and present alternatives in a path selection task. In the moderately complex task we studied, we found this model benefited from machine learning methods to set the discount amount with this model. In the fifth section we present Focal Points theory (Schelling, 1963) that describes a low-level cognitive ability to pick prominent solutions in the absence of communication. We found this model significantly increased the accuracy of a prediction model in a problem where people had to coordinate without communication.

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 (Selten, 1998). 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 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 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.

We studied what decision models, AAT or others, were used to solve two types of problems – a relatively simple optimization problem and a complex negotiation problem. 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. Full details of our implementation can be found in our previously published work (Rosenfeld & Kraus., 2009; Rosenfeld & Kraus, 2011).

In addition to AAT, other strategies, bounded and strictly rational, were possible here. A clear optimal strategy existed within the implementation of the commodity search domain. In the settings that we experimented with the specific strategy was – 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. We also recognize that AAT is not the only possible bounded model possible within this domain. Following Gigerenzer and Goldstein's fast and frugal heuristics (Gigerenzer & Goldstein, 1996), we would expect people 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. In fact, our previous work did find that people typically used these AAT strategies instead of optimal or fast and frugal heuristics.

We also analyzed a previously presented negotiation domain (Lin, Kraus, Wilkenfeld, & Barry, 2008). We consider a negotiation session that takes place after a successful job interview between an employer and a job candidate. In this session both sides wish to formalize the hiring terms and conditions of the applicant: her **Salary**, **Job Description**, **Car Benefits**, **Pension benefits** and **Working hours**. In the problem setting considered, each side could pick from a list of possible values for each of the parameters. For example, the employee might ask for a salary of 20000 per month, with the job title of Project Manager, with a car, pension benefits, and working 8 hours, while the employer might counter with the same offer, but a salary of only 12000 per month and without the pension benefits. The goal of this study is to accurately predict what each side would offer. Here again, equilibrium strategies were possible based on strictly rational behavior. Following Gigerenzer and Goldstein's model of fast and frugal heuristics we would have expected that simple compromise heuristics could be 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. Nonetheless, we overall found that people create aspiration based strategies where they negotiate for specific issues in a specific order. For example, we found that negotiations first focused on the salary parameter and only then move on to other parameters such as pension or car benefits. We found that adding these aspirations explicitly as a parameter for the machine learning models to consider helped significantly improve the accuracy in predicting people's offers.

Hyperbolic Discounting

The theory of discounted utility describes how people show preference to immediate payoffs versus delayed ones. Many of us know that certain activities are unhealthy– smoking, eating non-healthy foods, and not exercising enough. However, we prefer these behaviors as they provide immediate pleasure, despite their long-term consequences. We consider two different models of discounting utility – hyperbolic and exponential discounting. While both are widely used, experiments have compared the two and shown that hyperbolic discounting is often more accurate in explaining human (and even animals') decisions (Dasgupta & Maskin, 2005; Chabris, Laibson,

& Schuldt, 2006; Deaton & Paxson, 1993). However, one key question within this theory is the rate at which people discount their utility. For example, most people are willing to take pills or vitamins to improve their health, e.g. accept small discounts, while fewer are willing to take drastic lifestyle changes. As was true with the AAT studies, here too alternative models were possible, specifically those based on strictly rational models and machine learning.

A framework that is used to study decision making over time under uncertainty is the multi-armed bandit problem that was first introduced by Robbins in (Robbins, 1952). It is similar to a traditional slot machine but generalizes the slot machine to have more than one arm. When pulled, each arm provides a reward drawn from a distribution associated to that specific arm. Initially, the gambler has no knowledge about the arms, but through repeated trials, he gathers information on each of the arms. During the game, the player must balance between exploitation, or choosing the arm which performed best until the current time, and exploration, or trying new or less pulled arms.

To compare decision making theories for the multi-armed bandit problem we introduce the following path selection problem: Every morning a *driver* has several roads to choose from, which all lead to her office. The travel time on each road varies due to traffic; however each road is associated with some average travel time. The *driver*'s goal is to minimize the overall travel time. We consider a *system* which knows the exact travel time every day and provides the *driver* with advice regarding which road to choose from. The *system* also has knowledge about traffic along the various routes, giving it information about estimated fuel consumption of each of the routes. We assume that this *system* is self-interested and its goal is to minimize the *driver*'s fuel consumption rather than her travel time. For example, the system may be a government body which is trying to minimize the impact of burning fossil fuels, and thus aims to promote less pollution even if this comes a cost to a longer commute time for the driver. The driver must decide whether she will accept the system's recommendation or not. As the driver is aware that the system is self-interested, it must evaluate if its advise is worth accepting. However, on the other hand, the system has more information than the driver, and the driver might gain from listening to its advice.

In order for the system to better interact with the drivers, it is necessary to accurate model what types of advice are likely to be accepted. Towards this goal, we considered five different methods. The first method the strictly *rational* method is based on ϵ-greedy, which in known as a good method in multi armed bandit problems (Vermorel & Mohri, 2005). The multi-armed bandit problem was first introduced by Robbins in (Robbins, 1952) and is similar to a traditional slot machine but generalizes the slot machine to have more than one arm. When pulled, each arm provides a reward drawn from a distribution associated to that specific arm. Initially, the gambler has no knowledge about the arms, but through repeated trials, he gathers information on each of the arms. During the game, the player must balance between exploitation, or choosing the arm which performed best until the current time, and exploration, or trying new or less pulled arms. In this method we treat the advice that is generated by the system as another possible arm. If the driver chooses the advice he simply follows the road given by the advice. The prediction in this method is the road which has the highest chances to be chosen by ϵ-greedy.

At the other extreme, we considered two machine learning methods. The second method, *learning*, used the support vector machine (SVM) machine learning algorithm to learn which advise is accepted based on historical data of all other users' decisions. This data included the average time observed by the driver on each of the roads, the average time observed by the driver of when following the advice, and the actual number of times the driver chose each road and followed the advice. We also added information on that user's previous choice. We used 10 fold cross-validation to validate this model. The third method *sparse learning*, is similar to the *learning* method, but uses only 10% of the data and is tested on the remaining 90% of the date.

We also considered 3 types of psychological based models to predict people's decisions. *Exponential Smoothing*, *Short-term Memory* and *Hyperbolic Discount* are based on principles known from behavioral science and assume logit quantal response (Haile, Hortasu, & Kosenok, 2008). Quantal response suggests that instead of choosing the action with the highest expected utility, humans are known to choose actions proportionate to their expected utility (actions with higher expected utility are more likely to be chosen, but also actions with lesser utility have a positive probability to be chosen). Under the logit quantal response assumption, the probability for a person to choose action a' with utility $u(a')$ from a set of actions A is given by $p(a') = \frac{e^{\lambda u(a')}}{\sum_{a \in A} e^{\lambda u(a)}}$ where λ is some parameter (Haile et al., 2008). However, the value of this parameter is not clear, and must be learned from people's data. Exponential smoothing, or *ES*, is a method proposed by (Gans, Knox, & Croson, 2007) and is defined as follows. At $t = 0$ all actions start with some default value. Given $0 < \gamma < 1$, at each day t for a chosen action a, we let $ES_a(t) = \gamma \cdot \tau(a) + (1 - \gamma) \cdot ES_a(t - 1)$. An action that wasn't chosen maintains its previous value. This method is the equivalence of exponential discounting for discounting the past. Hyperbolic Discount, or *hyper* is a model that uses hyperbolic discounting of past actions. Formally, At $t = 0$ all action start with some default value. $hyper_a(t) = \sum_{t'<t} \frac{\tau_{t'}(a)}{f \cdot (t-t')}$ in case that $\tau_{t'}(a)$ is unknown for time t' (since a different road was chosen), $\tau_{t'}(a)$ is replaced by the default value. f is a parameter depicts the discount factor. The *Short-term Memory* model assumes that people have short memory and any instances previous to a “magic number” of the past 7 events do not influence their decisions. For more information about short-term memory see (Miller, 1956). All three psychological based methods attempted to learn all parameters with only 10% of the original data.

Focal Points

Focal points were introduced by Schelling in (Schelling, 1963) as a prominent subset of solutions for *tacit coordination games*, which are coordination games where communication is not possible. In such games (also known as *matching games* in game theory terminology) the players only have to **agree** on a possible solution, regardless of the solution itself. In other words, they receive a reward by selecting the same solution, regardless of the solution. When their solutions differ, both players lose and do not get any reward. A solution is said to be “focal” (also “salient”, or “prominent”) when, despite similarity among many solutions, the players somehow converge to this solution.

A classic example of focal point coordination is the solution most people choose when asked to divide \$100 into two piles, of any size; they should attempt only to match the expected choice of some other, unseen player. More than 75% of the subjects in Schelling's experiments created two piles of \$50 each; that solution is what Schelling dubbed a focal point. Here again, other behavioral models are possible – using decision theory would result in a random selection among the 101 possible divisions, as the (straightforward) probability distribution is uniform.

Several attempts have been made to formalize focal points from a game theoretic, human interaction point of view ((Janssen, 1998) provides a good overview). However, that research does not provide the practical tools necessary for predicting people's actions. In a meta-analysis of previous focal points experiments we developed some general properties that "focalize" an answer: (1) Centrality, (2) Extremeness, (3) Firstness, and (4) Singularity. Briefly, described these properties are as follows: Centrality is a rule that gives prominence to choices directly in the center of the set of choices, either in the physical environment, or in the values of the choices. Extremeness gives prominence to choices that are extreme relative to other choices, either in the physical environment, or in the values of the choices. Firstness is the rule that gives prominence to choices that physically appear first in the set of choices. It can be either the option closest to the agent, or the first option in a list. Singularity is the rule that gives prominence to choices that are unique or distinguishable relative to other choices in the same set. For further details and examples, we encourage the reader to refer to our previous work (Zuckerman, Kraus, & Rosenschein, 2011).

The task of learning which of these properties will be used by people is far from trivial due to the large number of possibilities. In contrast, the learning task was much simpler in the path selection domain where the discount value needed to be learned, or the limited number of parameters which may aspired for in the negotiation domain. To overcome this difficulty, we present a *Focal Point Learning* approach which combine this psychological approach and machine learning. To accomplish this we preprocess raw domain data, and place it into a new representation space, based on the focal point properties. Given our domain's raw data O_i, we apply a transformation T, such that $N_j = T(O_i)$, where i, j are the number of properties before and after the transformation.

The new feature space N_j is created as follows: each $v \in O_i$ is a vector of size i representing a game instance in the domain (world description alongside its possible choices). The transformation T takes each vector v and creates a new vector $u \in N_j$, such that $j = 4\times$[number of choices]. T iterates over the possible choices encoded in v, and for each such choice computes four numerical values signifying the four focal point properties presented above. For example, given a coordination game encoded as a vector v that contains three choices (c_1, c_2, c_3), the transformation T creates a new vector $u = (c_1^c, c_1^e, c_1^f, c_1^s, c_2^c, c_2^e, c_2^f, c_2^s, c_3^c, c_3^e, c_3^f, c_3^s)$ of size 12 (3 possible choices $\times$ 4 focal point rules), where $c_l^{c/e/f/s}$ denotes the *centrality/extremeness/firstness/singularity* values for choice l. Note that j might be smaller than, equal to, or greater than i, depending on the domain and the number of rules used.

We designed a simple and intuitive tacit coordination game that represents a simplified version of a domain where an agent and a human partner need to agree on a possible meeting place. The game, coined "Pick the Pile" is played on a 5-by-5 square grid. Each square of the grid can be empty, or can contain either a pile of money or the game agents Each square in the game board is colored white, yellow, or red. The players were instructed to pick the one pile of money from the three identical piles, that most other players, playing exactly the same game, would pick. The players were told that the agents can make horizontal and vertical moves.

Experimental Results

In this section we present a survey of previously and new results that demonstrate when and how machine learning techniques can benefit from behavioral theories. In general, we found that in the relatively simple optimization problem, strictly rational, AAT models and machine learning converged on nearly identical results. In the more complex path selection domain, the discount rate was unclear within the hyperbolic model and machine learning methods were able to learn the best value for this parameter. This combined model was more successful than an SVM machine learning model or other models based on strictly rational behavior. In the more complicated negotiation domain, adding information about people's aspirations increased the predictive accuracy of models built based upon machine learning. Strictly rational models performed far worse. In an even more complex coordination without communication domain, focal point information again improved the accuracy of a model based upon machine learning models. Strictly rational models and models built upon focal points without machine learning performed far worse.

Results from an Optimization Problem

In the first task, a relatively simple optimization problem, we wished to predict if a person would stop their commodity search in any given store. In this domain an optimal search strategy exist, namely, in the specific settings that we considered, the person should stop the search in the first store with a price less than 789. Note that this solution can be mathematically calculated and does not require any input from observed behavior. At the other extreme, we can create a prediction model exclusively based on machine learning techniques. Previously, we used decision trees to create this model. The advantage to specifically using this type of models lies in the output – we can check if the decision tree's decision model is consistent with the optimal solution or with other bounded models. We considered two such bounded models: fast and frugal heuristics and AAT. According to the more simple fast and frugal heuristic approach we would expect people to stop their search based on only one parameter, such as the number of stores visited to date, or the price of the commodity in any given store. According to AAT we would expect to see more complicated strategies with multiple parameters and some type of ordering and retreat between them. Our previous work (Rosenfeld & Kraus, 2011) did in fact find that the decision trees output was consistent with AAT strategies as people typically would immediately buy the commodity if it was below a certain price, but settle on a higher price after visiting a certain number of stores.

In this paper, we focus on when and how we can combine various decision theories to better predict people's decisions. In this domain, this included comparing the following models: 1. An optimal model based on expected utility – e.g. people buy only if the price is less than 789. 2. A machine learning model based on observed decisions. 3. A combination model. In this problem, the combination model involved adding information about the average price where people stopped their search, and the average number of stores after which they were willing to settle on a more expensive commodity. Note that here, as well as in all of the domains we consider, this hybrid approach assumes that we have some general information about a given population.

For this domain, we found that adding general information about people's aspirations was useful, but only slightly. Figure 1 presents the accuracy of different models in predicting when 41 people stopped their commodity search. Each of these people was presented with a simulation of the commodity search domain and ran at least 25 simulations where they eventually bought the commodity, logging a total of nearly 5000 instances where these people either decided to buy the commodity or to continue their search. The first column of Figure 1 presents a baseline *Naive* model that classifies all decisions based on the majority class, here assuming people will always continue the search. In the second column, we present the predictive ability of the optimal model – 82.8%. Column 3 presents the results from the machine learning method which performed similarly at 82.67% accuracy. Adding information from people's aspirations did help, but only slightly, with a 83.45% accuracy achieved through knowing the average values of these people's aspirations. Note that this value serves as an upper baseline, as we collected this aspiration data from the same population being evaluated. A more realistic aspiration model is the Sparse AAT model which used only 50 randomly selected decision to help model people's decision (or less than 1% of the total logged data). Nonetheless, even this model did slightly outperform both the optimal and based machine learning methods with an accuracy of 83%. This result is even more striking when you consider that machine learning models were validated through cross-validation of 90% of the data used for training the model, while this sparse model used less than 1% of the data. Thus, we conclude that in this relatively basic domain, differences between the predictive abilities of the different models was not large. Nonetheless, a slight improvement in prediction accuracy was obtained through limited information about people's aspirations.

Path Selection Results

Recall that the goal of the path selection domain is to predict what a person will do when receiving advise from a self-interested system generates. We generated three different types of advise ranging from fully self-interested to self-less. Subjects receiving the first type of advice were always advised to choose the road which was best for them, or the road that was the least time consuming. The second advice method always advised the subjects to choose the road which was best for the system, or the road that used the least fuel. The third advice tried to minimize some linear combination of both the fuel and time consumption. Each person received only one type of advise, but was unaware about which type of advice he was receiving. We intentionally used results combined from three different types of advice in order to build an accurate model of human behavior which will be true for a broad variety of advices. We performed trials with nearly 75 people – 22 were in the first group, 24 subjects in the second group and 24 in the third group. Each subject played 25 interactions. Results are shown in Table 2.

From the results we notice that people do not try to maximize their expected monitory value, and ϵ-greedy performs badly in predicting human behavior with only 45% of accurate predictions. Using the Support Vector Machine (SVM) machine learning methods on the data raises the prediction to 61.14%, and even using a limited training sample of only 10% of the data yields a prediction accuracy of 57.4% with this learning algorithm. All three psychological based models, which use only 10% of the data for learning, reach significantly better results ($p < 0.01$) when compared with the machine learning model with sparse data. Significant test was performed using binomial test. Although the short memory model is at par with the machine learning model with the full data set, both the *ES* and the *hyper* methods perform significantly better than the machine learning with the full data set (*SVM*). when comparing *ES* with *SVM* we reach $p < 0.05$ and when comparing *hyper* with *SVM* we reach $p < 0.001$.

AAT in a Negotiation Domain

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. Our premise is that as the negotiation domain is more complex than the optimization problem, one should add people's aspiration information into traditional models such as C4.5 to more accurately predict what bids people will offer.

To test this hypothesis, we proceeded to study what gain, if any, did adding AAT information have in predicting how people will negotiate. In the problem we considered, the parameters to be negotiated could have between 2 and 4 discrete values. In order to study this point we considered several models for the negotiation problem (see Table 3). The goal of all of these models was to predict the next value for each parameter. 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**. These strategies are based on previous work in these problems (Lin et al., 2008). However, as the equilibrium strategy depends 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.

Table 3 demonstrates the effectiveness of adding AAT information to boost prediction accuracy. The first row of this table 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 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 the significant improvement gained from the AAT information is significant and only one parameter did not gain from the added aspiration information. In this parameter, few instances existed where people had clear general aspiration changes, preventing any accuracy boost from this approach. Finally, the last line in the table 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.

Experimental Results for Focal Points in the Pick the Pile domain

In order to evaluate the effectiveness in adding focal point information in predicting people's actions, we conducted the following experiment. We collected data using an Internet website which allowed players from all over the world to participate in the game, and their answers were recorded. Each game session was constructed of 10 randomly generated instances of the domain. The *call for players* was published in various AI related forums and mailing lists all over the world, and eventually we gathered approximately 3000 game instances from over 275 different users from around the world.

We then compared the correct classification performance of both C4.5 learning trees and FFBP neural network classifiers. The comparison was between a **domain data agent** — an agent that was trained only on the raw domain encoding, a **focal point agent** (FP) — an untrained agent that used only the focal point rules for prediction, weighted uniformly, and a **focal point learning agent** (FPL) — as described above. "Correct classification" means that the agent made the same choice as that of the particular human player who played the same game. Obviously the learning problem is extremely difficult as there is no simple function that can capture the notion that for some games, different human players can select different choices.

We optimized our classifiers' performance by varying the network architecture and learning parameters, until attaining best results. We used a learning rate of 0.3, momentum rate of 0.2, 1 hidden layer, random initial weights, and no biases of any sort. Before each training procedure, the data set was randomly divided into a test and a training set (a standard 33.3%–66.6% division). Each instance of those sets contained the game description (either the binary or focal point encoding) and the human answer to it. The classification results using the neural network and the decision tree algorithms were very close (maximum difference of 3%).

Examining the results in Table 4, we see a significant improvement when using the focal point learning approach to train classifiers, rather than the domain data agent ($p < 0.01$ in two-proportion z-tests in all domains). In this domains, the domain data agent is not able to generalize sufficiently, thus achieving classification rates that are only about 5%–10% higher than a random guess (which is 33%). Using FPL, the classification rate improved to more than 65% correct classification. Since even humans do not have 100% success with one another in these games, FPL is correspondingly the more impressive. The results also show that even the classical FP agent, which does not employ any learning algorithm, performs better than the domain data agent, with 48% correct classification. In an additional analysis that was done on the FP agent, we saw a tendency in which the FP agent, when facing coordination problems with low *focality difference*, has its performance deteriorate to that of random guesses.

An additional advantage of using FPL is the reduction in training time (e.g., in the *Pick the Pile* domain we saw a reduction from 4 hours on the original data to 3 minutes), due to the reduction of input size. Moreover, the learning tree that was created using FPL was smaller, and can be easily converted to a rule-based system as part of the agent's design.

Conclusion

Predicting people's decisions is an important but complex task. To address this task, researchers often propose general behavior models such as rationality theory, or purely statistical methods such as machine learning algorithms. However, there often exist specialized cognitive models or theories that describe various tendencies or biases that are commonly used by the majority of the people. Such theories include bounded rationality theories, various risk attitudes, and use of heuristics.

This paper addresses how one can take a potentially relevant cognitive theory and use machine learning methods to help augment it to provide added value in predicting human behavior. We showed how three cognitive theories: Aspiration Adaptation theory, Hyperbolic Discounting theory, and the Focal Points theory could be used in conjunction with machine learning algorithms to create an improved classifier. Possibly equally significant is the result that strictly rational models, and even many specialized cognitive models, often do not accurately predict people's decisions. Our results also show some positive correlation between the complexity of the problem domain and the improvement in performance when augmenting the cognitive model. In all but the most simple task we considered, we found that "traditional" strictly rational models were often poor indication of how people will act. As problems' complexity increases, machine learning algorithms ability to predict people's decisions decreases, and the greater the benefit these algorithms gain from the information provided by the cognitive models. As we present a generalized approach for how to combine cognitive theories with machine learning algorithms, we expect this approach to be generally applicability to a variety of new domains as well.

References

Chabris, C., Laibson, D., & Schuldt, J. (2006). Intertemporal choice. *The New Palgrave Dictionary of Economics*.

Chalamish, M., Sarne, D., & Kraus, S. (2008). Programming agents as a means of capturing self-strategy. In *Aamas '08* (pp. 1161–1168).

Dasgupta, P., & Maskin, E. (2005). Uncertainty and hyperbolic discounting. *The American Economic Review*, *95*(4), 1290-1299.

Deaton, A., & Paxson, C. (1993, April). *Intertemporal choice and inequality* (Working Paper No. 4328). National Bureau of Economic Research.

Gans, N., Knox, G., & Croson, R. (2007, December). Simple models of discrete choice and their performance in bandit experiments. *Manufacturing & Service Operations Management*, *9*(4), 383-408.

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

Haile, P. A., Hortasu, A., & Kosenok, G. (2008). On the empirical content of quantal response equilibrium. *American Economic Review*, *98*(1), 180-200.

Janssen, M. C. W. (1998). Focal points. In P. Newman (Ed.), *The new palgrave of economics and the law* (pp. 150–155). MacMillan.

Kaelbling, L. P., Littman, M. L., & Cassandra, A. R. (1998, May). Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, *101*(1-2), 99–134.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291.

Keser, C., & Gardner, R. (1999). Strategic behavior of experienced subjects in a common pool resource game. *International Journal of Game Theory*, *28*(2), 241-252.

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

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

Manisterski, E., Lin, R., & Kraus, S. (2008). Understanding how people design trading agents over time. In *Aamas '08* (pp. 1593–1596).

Miller, G. A. (1956, March). The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychological review*, *63*(2), 81–97.

Mitchell, T. M. (1997). *Machine learning.* New York: McGraw-Hill.

Murakami, Y., Minami, K., Kawasoe, T., & Ishida, T. (2002). Multi-agent simulation for crisis management. In *Ieee workshop on knowledge media networking (kmn02).*

Murakami, Y., Sugimoto, Y., & Ishida, T. (2005). Modeling human behavior for virtual training systems. In *Aaai* (p. 127-132).

Nash, J. (1951, September). Non-Cooperative Games. *The Annals of Mathematics*, *54*(2), 286–295.

Neumann, J. V., & Morgenstern, O. (1944). *Theory of games and economic behavior.* Princeton University Press.

Robbins, H. (1952). Some aspects of the sequential design of experiments. *In Bulletin of the American Mathematical Society*, *55*, 527535.

Rosenfeld, A., & Kraus., S. (2009). Modeling agents through bounded rationality theories. In *IJCAI-09* (pp. 264–271).

Rosenfeld, A., & Kraus, S. (2011). Using aspiration adaptation theory to improve learning. In *Aamas* (p. 423-430).

Russell, S. J., & Norvig, P. (2003). *Artificial intelligence: A modern approach.* Prentice Hall.

Schelling, T. (1963). *The strategy of conflict.* New York: Oxford University Press.

Selten, R. (1998). Aspiration adaptation theory. *Journal of Mathematical Psychology*, *42*, 1910–214.

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

Selten, R., Mitzkewitz, M., & Uhlich, G. R. (1997, May). Duopoly strategies programmed by experienced players. *Econometrica*, *65*(3), 517-556.

Simon, H. A. (1957). *Models of man.* New York: John Wiley & Sons.

Vermorel, J., & Mohri, M. (2005). Multi-armed bandit algorithms and empirical evaluation. In J. Gama, R. Camacho, P. Brazdil, A. Jorge, & L. Torgo (Eds.), *Machine learning: Ecml 2005* (Vol. 3720, p. 437-448). Springer Berlin / Heidelberg.

Zuckerman, I., Kraus, S., & Rosenschein, J. S. (2011). Using focal point learning to improve human-machine tacit coordination. *Autonomous Agents and Multi-Agent Systems*, *22*(2), 289-316.

Table 1

Comparing the Prediction Accuracy between Optimal, Machine Learning and AAT Based Models

Naive	Optimal	Learning	Learning + Complete AAT	Sparse AAT
78.56	82.8	82.67	83.45	83

Table 2

Prediction Rate for the Path Selection Problem

model	prediction rate
Rational	45%
Learning	61.14%
Sparse Learning	57.4%
Short Memory	60.95%
Exponential Smoothing	63.56%
Hyperbolic Discount	65.26%

Table 3

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

	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 1	44.4444	67.5926	69.4444	66.6667	41.6667	67.5926	59.568
Equilibrium 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	**62.963**	**68.5185**	**75.9259**	**71.2963**	**91.6667**	**76.8519**	**74.53705**
C4.5 + Complete	95.3704	89.814	100	96.2963	100	96.2963	96.2962

Table 4

Results from "Pick the Pile" domain

Random guess	Raw Encoding	Only Focal Point Rules	Focal Point Learning
33%	40%	48%	65%