Opponent Models and Heuristic Strategies for Simple Games

Timothy J. Vickery (vickery@wjh.harvard.edu)

Department of Psychology, 33 Kirkland St.

Cambridge, MA 02138

Abstract

This study introduces a model that describes the reasoning strategies of a population of players in simultaneous-move, one-shot games. In the past, models of such behavior have explicitly employed the concept of Nash equilibrium in players' models of other players. In this model, behavior is accounted for in terms of simple heuristic strategies and opponent modeling, rather than by recourse to the concept of Nash equilibrium. The model represents six types of boundedly rational players: three types who employ no model of their opponents, two types who model their opponents as employing simple heuristics, and one type who models the population as a mixture of different types of players. Results show promise for eliminating the concept of Nash equilibrium from players' models of other players. In pursuit of this goal, a new type of graphical model based on Influence Diagrams is developed. Uncertain Decision Diagrams are suitable for modeling human decision-making with respect to explicit mental models that include noisy estimates of utility, and can be extended to model players' models of other players.

Introduction

Agents often confront situations in which not only their personal choices, but also the choices of others, affect the subjective value of an outcome. Some examples of these situations are the negotiation of hunting rights between two tribes, the barter of goods at a market, and a game of rock, paper, or scissors. Such problems are of great interest to cognitive science, because humans evolved and exist in an environment replete with demands to compete and opportunities for cooperation. An understanding of how humans reason about competitive scenarios would improve our understanding of human behavior in numerous ways, allowing us to probe questions about the representation of utility, other agents, and the world, and enabling the construction of artificial agents that exploit or enhance the strengths and weaknesses of human decision making.

The model sketched here describes how humans make decisions in games, or formalized incentive structures involving multiple agents. We model reasoning strategies employed when a person has no reward history with a game (i.e., the games are one-shot, not repeated). The model describes a population of players as a mixture of player types, some of whom employ simple heuristic strategies, and others of whom employ simple models of opponents who play by these heuristic strategies.

Game Theory Background

Traditional game theory formalizes the basic problem of determining the optimal solution for all agents in a setting in which outcomes for agents are determined by joint decisions. Nash (1950) determined that any such scenario has at least one solution strategy from which no player can benefit by deviating. This solution, called Nash equilibrium, is an optimal solution to a game when all agents behave "rationally," in terms of decision theory. The games typically submitted to game theoretic analyses are analytically tractable: far simpler than games like chess, but useful for probing the capabilities of agents.

Two games employed by Stahl and Wilson (1995) to study human behavior are shown in table 1. The model presented in this paper will later be applied to the data they collected as humans played these games, which are fairly prototypical of situations studied in game theory. In these games, each subject is asked to examine each matrix, in turn, and choose T, M or B as a response. The player's reward is determined by values in the chosen row. Each cell represents the value of an individual's choice given the choices of her opponent, which correspond to the columns. Thus, if a player chose T and her opponent chose "M," she would receive the payoff in cell (T,"M"). The games used here are symmetric, meaning that the opponent's payoff table is the same.

The prescriptive qualities of game theoretic models are fascinating and useful as benchmarks to the success of algorithms in approximating solutions, but behavioral economists and psychologists have found that the traditional forms are unfortunately dysfunctional as descriptive models in many situations (cf. Camerer, 2003, who reviews many behavioral observations that disconfirm predictions relying on the concept of Nash equilibrium). The most successful models of human behavior in *repeated* games depict humans as employing algorithms that depend on reward history to approximate an optimal solution (for example, reinforcement algorithms). In general, humans do not seem to employ strategies that correspond to Nash equilibrium strategies either initially or over repeated interactions, but rather they often approximate the Nash equilibrium strategy in a manner consistent with reinforcement learning or similar algorithms that depend on historical factors.

Behavior matching the Nash equilibrium in one-shot (non-repeated) games would imply either infinitely recursive reasoning or explicit knowledge of how to calculate Nash equilibrium. Thus, there is little reason to believe that humans bring to a game any mechanism tuned to extracting an optimal strategy like the Nash equilibrium solution without any experience at all with a particular incentive structure. The behavioral data reported by Stahl and Wilson (1995) are testament to this: the Nash equilibrium strategy alone is entirely inadequate for explaining behavior. However, the authors presented a successful model of behavior in one-shot games, in which some players were modeled as having a concept of others

Table 1:Games (2/12) used by Stahl and Wilson (1995)

Game		"T"	"M"	"B"	Game		"T"	"M"	"B"
	T	25	30	100		T	75	40	45
1	M	40	45	65	**2**	M	70	15	100
	B	31	0	40		B	70	60	0

that included the Nash equilibrium strategy. This paper addresses whether or not a model of this sort can be successful without resorting to the Nash equilibrium concept in agents' models of others.

The discovery of consistent initial strategies in game solving has important consequences. For one, it makes choice behavior predictable in novel situations. It also has consequences for the course of learning. As repeated iterations of the same or similar situation occurs, the time course of learning is more predictable with such knowledge than by an algorithm based upon reward history alone. Knowing the initial states of learners makes prediction of the eventual stable state reliable in models of behavior in games that have multiple equilibria.

Following Stahl and Wilson (1995), the model developed here assumes that humans are making optimal choices with respect to a noisy internal representation of the world, including representations of other agents. The model relies on an extension to Bayesian networks (Pearl, 1988) that allows the incorporation of beliefs about others into an explicit model.

Core Concepts of the Model

Two core concepts underlie the model presented here. The first core concept is that of a heuristic strategy. The second is depth of strategic reasoning. These concepts motivate the structure of the model.

Heuristic Strategies

A heuristic is a simple rule of thumb or "ad hoc" strategy. A person who is knowledgeable of an incentive structure but carries no preconceived notions of her opponents' states of mind might choose randomly, or they might use a simple rule. In fact, an informal test of these hypotheses gives support to the concept of a heuristic strategy. The twelve games used by Stahl and Wilson (1995) were presented to ten subjects. The subjects were shown only the payoffs to themselves. Subjects were asked to make no assumptions about what the opponents would receive as payoffs. Almost all of the subjects responded in a manner that was consistent with one of two heuristic strategies.

The first tendency was to choose the row in which the summed payoff was greatest. We will call this heuristic strategy the Maximum Sum (MS) heuristic. This strategy is rational if the opponent is assumed to choose at random. A second tendency was to choose the row with the highest minimum payoff. I will call this strategy the Greatest Minimum (GM) heuristic. This strategy is consistent with the notion that most humans are highly risk-averse, and is sensible if you want to maximize the least reward you could possibly receive. These two strategies together predicted >90% of the choices made in the twelve games used here. A third type of "level-zero" thinking will be included, the uniform random strategy (U), in which choices are made with equal likelihood. A player employing U might misunderstand the situation or lack motivation altogether.

Depth of Strategic Reasoning

The second core concept is the modeling of opponent strategies. If an opponent's strategy is known, then solving for one's best strategy is a decision theoretic problem. In most situations the opponent's strategy is unknown. One way of approaching a situation in which you have no expectations is to view it from the perspective of the opponent, and then respond with the best response to her expected choice. The problem with this is that it can be carried to an arbitrary depth. A player might model an opponent as behaving on the basis of a heuristic strategy, or she might regard her opponent as behaving on the basis of a best response to a heuristic strategy, and so on. Due to practical restrictions it is unlikely that an infinitely recursive strategy could be employed by humans in such situations. However, it is possible that people make assumptions about other opponents that allow them to reduce the complexity of the problem to a manageable form. Empirical studies support a shallow depth of initial reasoning (e.g., Hedden and Zhang, 2002).

Stahl and Wilson (1995) constructed a model that is closely resembled by the model presented in this paper, and which will serve as a point of comparison. They modeled a population of players using a mixture of player models, and found support for the idea that the population consisted of players who randomly selected responses (U), players who chose the best response to U, or BR(U)[1], players who chose the best response to the best response to U, or BR(BR(U)), players who behaved as if their opponents played the Nash equilibrium strategy (Naïve Nash, or NN), and "worldly" types who behaved as if the world were a mixture of the above types. They used their model to estimate the posterior likelihood of each player's type, and found very strong support that most players were acting consistently with the behavior of one of the predefined types.

Two aspects of their study led to this model. For one, their model was rather arbitrary and it is difficult to imagine a generalization to more complex scenarios. The model presented here adopts a graphical modeling approach, making it more extensible by allowing the addition of arbitrary variables and levels of reasoning. Secondly, this model eliminates the concept of the Nash equilibrium from players' models, and replaces it with heuristic types described above.

---

[1] If estimates of utility are noiseless, BR(U) is equivalent to MS, since it prescribes the choice of row with highest average utility.

The Structure of the Model

This model assumes, like Stahl and Wilson (1995), that the population is composed of players who approach each individual game with a consistent initial strategy. Each player type is modeled individually using a graphical model from which the probability of any particular strategy can be derived. These models are the components of a mixture model which is used to obtain estimates of the probability of the data given the full model.

Graphical Models and Extensions

Bayesian networks are a class of graphical models in which the causal structure of the world is codified into nodes, representing variables, and edges, representing dependencies amongst variables. Nodes can have various states, which are either known or unknown. For example, a node might represent rain on the 25th. Before the 25th, the state of the node is unknown, but various other factors influence the probability of each state (e.g., the recent weather history). Given knowledge about the states of known variables, and the probabilistic dependencies amongst the nodes, the probability of an unknown variable's taking on some state can be determined. Bayesian networks have been explored extensively in computer science (Pearl, 1988), and recently applied to modeling human cognition. Influence diagrams (IDs) extend Bayesian Networks with decision and utility nodes (Howard and Matheson, 1984). They allow the modeler to calculate the best possible choice for an agent given specific knowledge about the world, the choices available to the agent, and the payoffs that depend upon the choices of the agent and the state of the world. Decision nodes in IDs represent choices of agents (represented as rectangles in diagrams), and simply denote the (discrete) choices available to the agent. Utility nodes represent the value to agents of states of the world. Decision and utility nodes possess many of the same properties of chance nodes in a Bayesian network. Many details are omitted here due to limited space.

Extensions of IDs called multiple agent influence diagrams (MAIDs) have been used recently to simplify the process of calculating the Nash equilibrium solution for players in competitive games (Koller and Milch, 2001). MAIDs allow modeling of multiple agents, each with their own decision and utility nodes. Solving a MAID involves calculating the optimal decisions of players with respect to their knowledge of the world, but MAIDs assume rational agents with infinitely recursive models of opponents. Given that the goal of this work is to eliminate the need for the Nash equilibrium concept in explaining human behavior, We will not spend any more time discussing the specifics of MAIDs.

An additional extension to MAIDs and IDs, known as networks of influence diagrams (NIDs), can be used to model agents who have any particular model of their opponents, agents who make decisions based on extraneous variables, and agents who conceive of a variety of possible models of their opponents, with uncertainty about which one their opponent will actually employ (Gal and Pfeffer, 2003). The models employed by this study borrow concepts from the NIDs framework, but the full arsenal provided by this modeling language is not necessary to build the basic models that we will use, and some additions are necessary.

NIDs are rooted, acyclic directed graphs in which the nodes (called blocks to avoid confusion with nodes internal to blocks) are self-contained IDs and MAIDs. Root blocks represent the "top-level" model. Blocks are assigned to individual agents. The decisions of agents may be modeled by child blocks. Edges from one block to another block indicate that a decision node in the parent is modeled by the child.

If a decision in a parent block is modeled by multiple child blocks, then a chance variable (labeled Mod[D], where D is the modeled decision node) is introduced to the parent. Mod[D] is a parent node D, and it takes on values corresponding to each of the child blocks that model the decision. Its conditional probability table (CPT) contains the probability allotted by the parent block to each of the different child models. In other words, it represents the degree of belief that each of the child blocks is being used by opponents. To solve a NID, one simply works from the leaf blocks to the root, solving for each decision variable the optimal response. Leaves are MAIDs or simple IDs, which may be solved according to known methods. The decision rules thus computed are available to the parent nodes, and are incorporated into parents as follows: each decision node in a parent that is modeled by a child requires the addition of a chance node to the parent. For each edge leaving a block, a chance node is added to the parent. This node has a conditional probability distribution over its component choices that is determined by the solution to the child. The original decision node becomes a chance node with each node that has been added as its parents. The new node takes on values of its parent nodes according to the probabilities allotted to the CPT of Mod[D]. Once the tree has been solved up to the root node, the root node becomes a MAID in the same fashion. Solving the root gives a Bayesian network that is open for interesting queries. This description is dense, high-level and misses many nuances. For a full description of NIDs, see Gal and Pfeffer (2003).

Traditionally, IDs have been used to model situations for the purposes of making a decision. However, since human decision making is inherently noisy, in order to model human behavior we will have to replace the traditional decision nodes with a noisy version. To accomplish this, each decision node is converted to a chance variable where each choice, or state, is chosen with some probability that is determined on the basis of a *noisy* estimate of expected utility. If the noise has the properties that it is additive, independent and identically distributed according to a Weibull distribution, then the
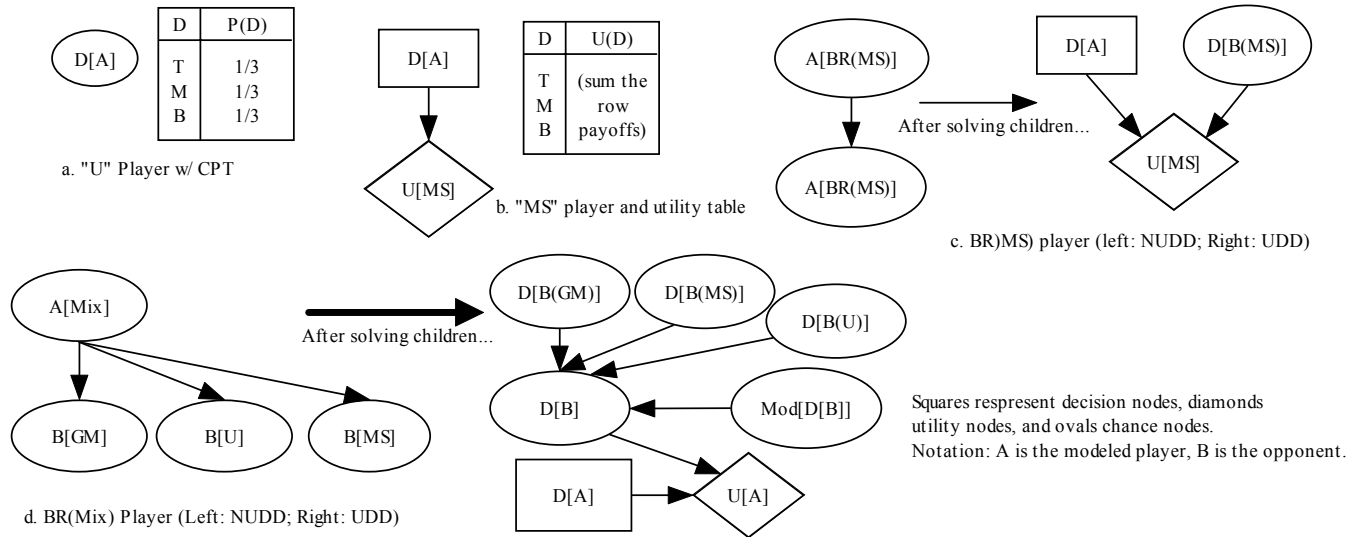
Figure 1: Four models of players in one-shot, three-choice matrix games

probability of the decision takes the convenient conditional logit form,

$$P_j(\gamma) \equiv \frac{\exp(\gamma * \mathrm{EU}_j)}{\sum_k \exp(\gamma * \mathrm{EU}_k)} \quad (1)$$

where P_j is the probability of making choice j, EU_j is the expected utility of the choice, and γ is a noise parameter (this choice was also inspired by Stahl and Wilson, 1995). Expected utility is calculated using basic algorithms for IDs, on the basis of the values in utility and chance nodes. IDs with the addition of this new type of decision node will be referred to as Uncertain Decision Diagrams (UDDs).

Another modification to NIDs for modeling humans is the elimination of decision nodes for decisions that are modeled by child blocks. The reason for this is that leaving these decision nodes in for the opponent would require the use of MAIDs. Algorithms for solving MAIDs introduce the Nash equilibrium concept, which we attempt to eliminate as an explanation for behavior in this game. Graphical models with the above modifications will be referred to as NUDDs (Networks of UDDs).

A model of behavior in one-shot matrix games

Using NUDDs, behavior in any simultaneous matrix game can be modeled given assumptions about the contents of utility nodes and the form of each player's model. Following Stahl and Wilson (1995), this model assumes that players fall into several categories. The primary goal was to replicate Stahl and Wilson (1995), but to do so without appealing to the Nash Equilibrium in opponent models. Their notation is reproduced where possible, and many of the same procedures are followed to maximize the comparability of this model to theirs.

Of the five player types posited by Stahl and Wilson, two include the concept of Nash equilibrium (i.e., some component of these two models specifies that the

opponent's prior over responses is the same as the Nash equilibrium solution to the game). Revisiting this model, we will define 6 types of our own, without appeal to the Nash equilibrium concept. Descriptions of these 6 types follow, along with the parameters that were allowed to vary for each model. For modeling simplicity, the values of these parameters were assumed to be constant for every player within a class.

1. The level-0 type of Stahl and Wilson (1995) becomes the "Uniform" player (U). A model of this player is pictured in figure 1a. Her choices are simply represented as a chance node with equal probabilities for each of the three decisions. This model has no free parameters.
2. The level-1 type (best response to level-0) becomes the "Maximum Sum" heuristic player (MS). This player's model is shown in figure 1b. The model is an UDD with a decision node and a utility node, where the utilities have been determined by summing the payoffs for each row. The noise parameter γ_1 was allowed to vary.
3. The level-2 type (best response to level-1), becomes the "Best response to maximum sum heuristic" player, BR(MS). The model for this player type is shown in figure 1c (in NUDD form and solved form). This is a simple NUDD in which the child block is the UDD described in 1. The noise parameter γ_2 was varied.
4. A second type of heuristic player, who chooses on the basis of the row with the maximum minimum payoff. This player is referred to as the "Greatest Minimum" heuristic player, or GM. Her model is just like the MS player, but the values of the utility node correspond to the least minimum payoff for each of the choices. Noise parameter, γ_3, was allowed to vary.
5. The fifth type is the "Best Response to Greatest Minimum Heuristic" player, or BR(GM). This model is similar to BR(MS) with the child block replaced

with the GM player's UDD. Decisional precision γ_4 was allowed to vary.

6. The sixth type is the "Mixture of strategies" player, or BR(Mix). This is a player who believes that others are a mixture of the above strategies. We include a player who chooses the best response to a mixture of strategies 1, 2, and 4. In other words, she models her opponents as playing some mixture of the uniform and heuristic strategies. This player has 3 parameters that are allowed to vary: γ_5 (noise), μ_1 (the proportion of her opponents that she believes are uniform players), and μ_2 (the proportion believed to be MS players). This player believes that opponents are GM players with probability $\mu_3 = 1 - \mu_1 - \mu_2$. This player's model is shown in figure 1c. The CPT for Mod[D] is determined by the mixing parameters described above.

Each of the models above produces predictions of the probability of each choice in each game. Choice probabilities for each strategy $j \in \{1, 2, 3\}$ within each game $i \in \{1, 2, ..., 12\}$ for each model l (l=1, 2, 3, 4, 5, or 6) were determined by equation 1, where expected utility of each response is calculated using a common algorithm for solving influence diagrams based on the structure of the graphical model. Subject h's strategy in game i will be referred to as $s(h, i) \in \{1, 2, 3\}$. The probability of choice i given model type l and model parameters β_l is denoted $P_{ij}(\beta_l)$. Following Stahl and Wilson's (1995) conventions, the probability of subject h's choices in game i conditioned on that player belonging to class l is $P_{is(h,i)}(\beta_l)$ and the joint probability of all of participant h's choices conditional on being a class l player is

$$P_l^h(\beta_l) = \prod_i P_{is(h,i)}(\beta_l) \qquad (2)$$

The full mixture model includes 6 parameters α_l defining the proportion of subjects in the population who employ model l (the sum of all α_l is 1). The likelihood of participant h making her particular choices is given by

$$L(s^h \mid \beta) \equiv \sum_{l=1}^{6} \alpha_l P_l^h(\beta_l) \qquad (3)$$

Log-likelihood of the entire sample is

$$\mathcal{L} \equiv \sum_h \log[L(s^h \mid \beta)] \qquad (4)$$

A Test of the Model

Parameter estimates β' were determined by maximizing (4) for the data collected by Stahl and Wilson (1995). They tested 48 subjects on 12 games (2 of which are shown in Table 1). Estimates of the maximum-likelihood parameters were obtained using the constrained line-search procedure provided with the Matlab Optimization Toolbox and several randomly chosen starting points. Imitating the procedures of Stahl and Wilson (1995), we constructed 95% confidence intervals using a bootstrap technique. From the parameter estimates β', M=400 sample decisions s^* (of the same sample size, 48 subjects and 12 games) were generated using the following technique. First, draw a uniform random deviate [0,1] and choose a class based on this value and the $\hat{\alpha}_l$'s (i.e., choose class l with probability $\hat{\alpha}_l$). Then, obtain the probability of each decision given the corresponding model and its parameter estimates. Finally, choose a decision by drawing a second uniform random deviate and comparing it to this value. Once the samples s^* were generated, we determined maximum-likelihood parameters β'^* for each s^* using the same technique as before, and determined 95% confidence intervals for each β. The maximum-likelihood estimates and confidence intervals are given in Table 2. On first glance, we see that most parameters take on reasonable values. Approximate proportions of each player type were 14% (U), 17% (MS), 2% (BR(MS)), 35% (GM), 17% (BR(GM)), and 16% (BR(Mix)). Their confidence intervals do not include 0, suggesting that sufficient evidence exists to include all of these types in this model, although two of the intervals do come close. The only immediate cause of concern is the confidence interval of γ_2, which includes 0. When γ_2 is 0, noise is so great that the strategy becomes the same as U. The likely reason for this is that only one player (as will be shown) is employing this strategy, and thus in many fits, γ_2 was allowed to vary greatly with little consequence to the likelihood values. The BR(Mix) strategist seems to be employing a model with 40% GM players, 20% MS players, and 40% U players. The maximum log-likelihood was -448.75, which is slightly less than the value of -442.73 obtained by Stahl and Wilson (1995). However, we believe that it still compares favorably, especially given the following considerations: (1) although this model's parameters have 12 degrees of freedom relative to their 11 degrees of freedom, our model is somewhat more restricted due to the replacement of the Nash equilibrium prior. The Nash equilibrium prior is more flexible than either of our heuristic priors, predicting in some cases equal probabilities for all three choices. (2) Goodness-of-fit is lower for our (conceptually simpler) model. This statistic was obtained using the following equation,

$$\lambda = \sum_i \sum_j \frac{(n_{ij} - N\pi_{ij})^2}{N\pi_{ij}} \qquad (5)$$

where n_{ij} is the number of people who chose j in game i, N is the number of subjects, and π_{ij} is the proportion of such choices predicted by our model. The π_{ij} are calculated as follows,

$$\pi_{ij} = \sum_{l=1}^{6} \hat{\alpha}_l P_{ij}(\beta_l) \qquad (6)$$

We obtained λ=49.48 (distributed chi-square with df=24) for the full 12 games and 48 subjects, which exceeds acceptable values (p<.01). The null hypothesis that this model underlies the production of the data must be rejected on the basis of this statistic. However, this difficulty was also encountered by Stahl and Wilson (1995), who obtained an even higher value, λ=57.57. They point out that games 10 and 11 produced particularly troublesome results for their model, and the same applies here. Excluding these games from analysis, I

obtained λ=25.79 (df=20), compared with their λ=26.09. The model cannot be rejected (at α=.05) as an explanation of this data on the basis of this statistic. A direct comparison of these models is not possible, since they differ structurally. However, the above comparisons are not unfavorable to this model.

Table 2: Results of model fit to data from S. & W. (1995)

Parameter	Estimate	95% Confidence Interval	
$\gamma_{1[MS]}$	0.2718	0.1815	0.4583
$\gamma_{2\,[BR(MS)]}$	0.5549	0	3.9587
$\gamma_{3\,[GM]}$	0.0710	0.0504	0.0917
$\gamma_{4[BR(GM)]}$	0.2903	0.1678	0.4174
$\gamma_{5[BR(Mix)]}$	0.5525	0.3630	1.0630
$\mu_{2\,[\%\,MS\,in\,BR(Mix)]}$	0.1989	0.0992	0.2619
$\mu_{3\,[\%\,U\,in\,BR(Mix)]}$	0.4016	0.3077	0.5517
$\alpha_0\,[U]$	0.1691	0.0478	0.2918
$\alpha_1\,[MS]$	0.0208	0.0003	0.1490
$\alpha_2\,[BR(MS)]$	0.3450	0.2010	0.5552
$\alpha_3\,[GM]$	0.1658	0.0992	0.2619
$\alpha_4\,[BR(GM)]$	0.1591	0.3077	0.5517
$\mathcal{L}$	-448.75 (compared to -442.73)		
λ	49.48/25.79 (compared to 57.57/26.09)		

Posterior estimates of player types were obtained using a bootstrap method described by Stahl and Wilson (1995). Only a sketch of the results is reported here. Results show that subject classifications were not as well-defined as in the Stahl and Wilson (1995) model. 31 out of 48 were classified with a probability greater than 90%, compared with 38 in their model. However, many fewer bootstrap parameter estimates were obtained, and each of these was from a small number of attempts to search the parameter space. We also did not constrain parameter values as much as my predecessors. Even so, it is clear that the new heuristic (GM) and the model that responds to GM, BR(GM), are most appropriate for 18 and 10 players, respectively, and most of these with extremely high probability. At the very least, there is evidence for the employ of these strategies in lieu of NN. MS picks up 8 players (the same as their roughly equivalent BR(U). BR(MS), which in our model finds very little support, is still only likely for one subject. Given that one player's decisions match the noiseless BR(MS) predictions exactly, it is probable that this player used that strategy, but there is little support for that strategy, otherwise. Br(Mix) is most probable for only 4 players. The uniform strategy is the most likely candidate for only 7 players. The GM strategy is highly successful, with extremely high probability of being the model for many players formerly classified as level-0 (U) and NN players, as well as several mixture players. BR(GM) is also successful in explaining the activities of players formerly thought to be playing a best response to a mix of NN, U, and BR(U) strategies.

Conclusions

It is important to understand how real agents form their initial models, not only to predict agent behavior in novel situations, but also to predict the course of learning. This study shows that a successful model need only assume simple heuristics and low level opponent modeling to predict behavior. This model shows promise as an explanation for human behavior in simple games played without repetition. The model fit comparably to the model of Stahl and Wilson (1995), and most importantly it eliminated the Nash equilibrium concept from players' models. The heuristics proposed in this article require only very simple cognitive abilities.

A new type of node was introduced to Influence Diagrams to make them suitable for modeling human decision making. UDDs should be extended to model sequential decisions. Any domain that involves decision-making that employs internal models of other agents might benefit from the application of such models. In future work, the methods here will be extended to determine which models are most appropriate to include in a full mixture model. These models should be extended to model different types of games with more choices, more agents, and sequential decisions.

Acknowledgments

References

Camerer, C.F. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*, Princeton: Princeton University Press.

Gal, K. and Pfeffer, A. (2003). A Language for Modeling Agents' Decision Making Processes in Games. *Second International Joint Conference on Autonomous Agents and Multi-Agent Systems*, Melbourne Australia.

Hedden, T. and Zhang, J. (2002). What do you think I think you think? Theory of mind and strategic reasoning in matrix games. *Cognition*, 85: 1-36.

Howard, R.A., and Matheson, J.H. (1984). Influence Diagrams. In *Readings on the Principles and Applications of Decision Analysis*, 721-762.

Koller, D. and Milch, B. (2003), Multi-Agent Influence Diagrams for Representing and Solving Games, *Games and Economic Behavior, 45*: 181-221.

Nash, J.F. (1950). Equilibrium points in n-person games. *PNAS, 36*:48-49.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems,* San Francisco: Morgan Kaufmann Publishers.

Stahl, D.O., and Wilson, P.W. (1995). On Players' Models of Other Players: Theory and Experimental Evidence. *Games and Economic Behavior*, 10, 218-254.