In Proceedings of the Sixth International Conference on Cognitive Modeling, 255-260.
Mahwah, NJ: Lawrence Earlbaum.

Cognitive modeling versus game theory: Why cognition matters

Matthew F. Rutledge-Taylor (mrtaylo2@connect.carleton.ca)
Institute of Cognitive Science, Carleton University, 1125 Colonel By Drive
Ottawa, Ontario, K1S 5B6 Canada

Robert L. West (robert_west@carleton.ca)
Institute of Cognitive Science, Department of Psychology, Carleton University, 1125 Colonel By Drive
Ottawa, Ontario, K1S 5B6 Canada

Abstract

We call into question game theory, as a account of how people play two player zero-sum games. Evidence from a modified version of the game Paper, Rock, Scissors suggests that people do not play randomly, and not according to certain play probabilities. We investigated the relationship between game theory predictions and a cognitive model of game playing based on the detection of sequential dependencies. Previous research has shown that the sequential dependency model can account for a number of empirical findings that game theory cannot. The sequential dependency model has been implemented using both simple neural networks and ACT-R. In this paper we used simple neural networks (a description of how our findings relate to the ACT-R model is included in the Conclusion section). For simple games, such as Paper, Rock, Scissors, game theory has been able to correctly predict aggregate move probabilities. In this paper we show that this is an artifact of the symmetry of the payoffs, and that for asymmetrical payoffs the game theory solution does not predict human behavior. Furthermore, we show that the model of game playing that underlies game theory cannot be used to predict the results no matter what move probabilities are used. Finally, we show that the results can be accounted for by augmenting the network sequential dependency model so that the reward system is related to the game payoffs.

Introduction

Game theory (VonNeumann & Morgenstern, 1944) was not created as a cognitive model, that is, it was not intended to account for mechanisms underlying the manner in which humans play games. Rather it was intended to be a formal mathematical system for understanding game playing from a rational perspective. However, game theory has had an enormous influence on theories and ideas about how people play games. Because empirical studies have shown that game theory is poor at predicting how humans play games, most researchers do not believe that game theory is a good model of how humans process information during games (Pool, 1995). Nevertheless, the game theory model of a player is still extremely influential in terms of how we understand human game playing.

Game theory is a large collection of mathematical principles that are used in accounts of a variety of human interactions. It is used by a variety of researchers including economists, mathematicians, and behavioral psychologists to name a few. The focus of this paper is specific to the application of game theory as an account how people play a relatively constrained set of non-zero sum games. Henceforth, the use of the term 'game theory', will be used to refer to a subset of principles within game theory commonly used to describe this type of game.

In game theory, a game is defined as a situation in which two players each select one move from a discrete set of choices, the combination of which determines a payoff for each player. There are two varieties of games. In zero-sum games the payoffs for the players must sum to zero, and there is no option for cooperation (i.e., one player's gain is always equal to the other player's loss). In non-zero-sum games this constraint does not exist and cooperation can be an option. In either case, to apply game theory it is necessary to assume that players can make random selections from their move choices according to specific probabilities assigned to each move. This amounts to two assumptions about the cognitive abilities of players: (1) they have some way of calculating or learning move probabilities, and (2) they are able to select moves at random according to these probabilities. In what follows, we will refer to these assumptions concerning players' abilities as *the game theory player model*. Assuming that a player matches the game theory player model, it is generally possible to use game theory to calculate the optimal set of move probabilities. We will refer to the outcome of game theory calculations as *the game theory solution*.

Most research on human game playing has focused on non-zero-sum games, where it has been shown that people generally do not follow the game theory solution. One reason for this is that people are often influenced by factors that are extraneous to the game theory solution, such as social norms concerning cooperation (Poundstone, 1992; Samuelson, 1997). However, these results show only that people do not play according to the game theory solution. The game theory player model is still viable as it is possible to explain the results in terms of people playing according to the game theory player model (i.e., probabilistically), but with move probabilities inconsistent with the calculated game theory solution. In zero-sum games there is no option to cooperate so there should be less extraneous influences. However, there is very little direct research on human zero-sum game playing (Poundstone, 1992). The general view of human zero-sum game playing is based mainly on experimental comparisons between humans and the game theory player model. With regard to this, psychological studies have shown that people are very poor at the two essential skills required by the game theory player model: (1) learning optimal move probabilities (e.g., Gazzaniga, 1998), and (2) behaving randomly (see Wagenaar, 1972 for a review).

Based on this, a common view of individual humans as game players is that we are poor game theory players. That is, we play in a way consistent with the game theory player model but we are not good at learning the optimal move probabilities, we are poor at being random, and we are influenced by considerations extraneous to the game theory framework (e.g., norms about cooperation). Attempts have been made to adapt the game theory player model to better capture human behavior. The focus in this area has been on introducing a learning component to explain how humans acquire move probabilities. Game theorists have studied the effects of using learning algorithms to acquire move probabilities (Fudenburg & Levine, 1998), and there have been some attempts to build cognitively plausible versions of the game theory player model using cognitive architectures (e.g., see Ritter & Wallach, 1998 for examples using ACT-R and SOAR). However, all of this work is based on the implicit acceptance of the game theory player model as an appropriate framework for understanding human game playing behavior.

An alternative player model

A psychologically plausible alternative to the game theory player model is that instead of trying to learn advantageous move probabilities, people try to detect sequential dependencies in their opponents' play and use this to predict their opponents' moves (Lebiere & West, 1999; West, 1998; West & Lebiere, 2001). That is, players learn recognize patterns of consecutive moves in their opponents' play. This model is consistent with a large amount of psychological research showing that when sequential dependencies exist, people can often detect and exploit them (e.g., Estes, 1972; Restle, 1966; Rose & Vitz, 1966; Vitz & Todd, 1967). It also explains why people tend to do poorly on tasks that are truly random - they persist in trying to predict the outcomes even though doing so results in sub-optimal results (e.g., Gazzaniga, 1998; Ward, 1973; Ward, Livingston, & Li, 1988).

West and Lebiere (2001) used neural networks to examine the possibility that people play games by attempting to detect and exploit sequential dependencies in their opponent's play. The networks were designed to detect sequential dependencies in the game of Paper, Rock, Scissors (hence forth PRS). PRS was chosen because it is familiar to most people and because it is very easy to play. It is also a zero-sum game and therefore does not involve the complications associated with the option to cooperate. The players were modeled using very simple two layer neural networks rewarded by adding 1 and punished by subtracting 1 from the connection weights (all of which started with a weight of 0). The inputs to the network were the opponent's previous moves (referred to as lags), and the outputs were the moves the player would make on the current play. The goal in creating these networks was to use the simplest possible model of sequential dependency detection.

The simulations revealed that processing more lags is an advantage. That is, a network that processed the last two lags (a lag 2 network) would reliably win against a network that processed only the last lag (a lag 1 network). Also a network that treated ties as losses (an aggressive network) could reliably win against a network that was neither punished nor rewarded for ties (a passive network). Furthermore, these effects were additive and approximately equal in magnitude. Another important finding was that the interaction between the networks produced a chaos-like behavior that made them appear to be playing randomly. Subsequent to examining the play of the neural network models, West & Lebiere (2001) investigated the play of humans against the models. They found that humans could reliably beat both the aggressive lag 1 network and the passive lag 2 network. This suggested that humans play similarly to the aggressive lag 2 network. Although there was a small but statistically significant tendency for people to lose against the aggressive lag 2 model rather than tie, this was attributed to the humans being unable to play as consistently as the network model. This interpretation was supported by the fact that subjects reported getting frustrated when playing the aggressive lag 2 network (i.e., playing hundreds of trials is only fun if you are winning). Both the network model and the game theory solution predicted that, on average, people would play each of the three play options with equal frequency. However, the game theory model (i.e., the combination of the game theory player model and the game theory solution) predicted that people would tie against the networks, which was not the case.

Aggregate behavior

The West & Lebiere (2001) results show that the sequential dependency player model can account for results in simple zero-sum games that game theory model cannot. In addition, the sequential dependency player model is consistent with the empirical facts. Specifically, people are poor at being random and poor at learning optimal move probabilities because they are instead trying to detect and exploit sequential dependencies. However, the game theory solution did correctly predict that, on average, humans played paper, rock, and scissors with approximately equal frequency. This raises the question of whether game theory can still be considered viable for predicting aggregate move probabilities for this type of game. That is, regardless of the details of how people play, does game theory capture certain higher-level stochastic properties of game playing behavior? After all, even if people do not process game information in the manner suggested by the game theory player model, it may still be the case that across time and across individuals, human game playing can legitimately be viewed as (pseudo) randomly emitting moves according to certain probabilities. To test this possibility and to probe deeper into the relationship between game theory and the sequential dependency player model we tested a variant of PRS.

Rock=2

An aspect of PRS that makes it a very simple game is that each of the three play options is functionally identical to the other two. That is, each move beats one of the other two moves and loses to the remaining move. In addition, a win is worth the same amount for each move. Thus, it is not surprising that the game theory solution for playing PRS is to play the three options with equal probabilities. The reason this is somewhat problematic is that the agreement between the game theory solution and human behavior for this game may be an artifact of the simplicity and symmetry of the game. To clarify this issue, a modified version of PRS was developed. The new game was identical to the original PRS game except that a win using rock counted for 2 points while a win with scissors or paper counted for only 1 point. In this way, each of the three choices were unique. Rock could win two points and lose only one, scissors could lose two points and win only one, and paper could win or lose only one point. The game theory solution to this modified version of Paper, Rock, Scissors differs depending on whether a zero-sum or non-zero-sum interpretation of the game is adopted. That is, whether a player is trying to maximize the difference in points between himself and the other player, or whether each player is attempting only to maximize the total number of points for themselves. However, since we instructed our subjects to try to maximize the points difference we will focus on the zero-sum interpretation. In this case, the game theory solution is to play paper 50% of the time, rock 25% of the time, and scissors 25% of the time, for the expected outcome of a tie.

The simulated opponents

For this study we used the same simple network models as West & Lebiere (2001) and created three simulated opponents for our human subjects to play against. The first two opponents were taken directly from the West & Lebiere (2001) study. They were, the aggressive lag 2 model and the aggressive lag 1 model. We did this to test the hypothesis that people simply try to maximize wins in this type of game. If this were the case then the results against these two models would replicate the results of West & Lebiere (2001) as neither the humans nor the models would be influenced by rock wins being worth more points. To create the third simulated opponent we adapted the aggressive lag 1 model so that it rewarded the relevant connection weights by 2 instead of 1 when it won with rock. This model was created to pit the human players against a model that might better take advantage of winning with rock, but still had some weaknesses that humans could exploit (i.e., it was only a lag 1 network and it was not set to avoid losing with scissors thereby allowing the opponent to win with rock). The reason for this was that games where humans win are much more informative than games where humans lose, as the loss can be attributed to extraneous factors such as lack of effort, boredom, or frustration. To distinguish this model we will refer to it as the rock=2 lag 1 model.

Method

Ten human subjects played against each of the three network models. They were instructed to try to maximize the points difference in their favor by as much as possible. This goal corresponded to the zero-sum interpretation of the game. The subjects were also told that the network models did not play randomly and that they could be beaten. Additionally, the subjects were instructed to play naturally, not to play too slowly, nor to think too much about their play. The order in which the subjects played the network models was random. Subjects were required to play one game (300 trials) against each of the three different network models. This process took from between 30 to 45 minutes in total depending on the speed at which the subject played.

Results and Discussion

The points differences between each of the human players and each of the network models were calculated for each trial by subtracting the network score from the human score. Thus a positive score indicated that the human was ahead and a negative score indicated that the network was ahead. The mean total points difference at the end of each game (see Table 1) revealed that the humans were able to win against all of the network models. To test the significance of this we ran a regression on the group points difference data for each different type of opponent across trials. The regression coefficients thus corresponded to the average rate of points accumulation (i.e., points difference/trials) for the humans against each network opponent (the intercept was forced through zero). 95% confidence intervals for the coefficient values revealed that all of them were significantly above zero. That is, against each network models, there was a significant tendency for the humans to win.

The fact that people could beat the aggressive lag 2 model under these conditions, whereas they tended to lose in West & Lebiere (2001), where all three varieties of wins were of equal value, indicates that they were able to exploit the fact they knew that wins using rock were rewarded with 2 points. Thus, the hypothesis that people simply try to maximize the number of wins regardless of the number of points awarded for wins, was refuted. That is, the profiles of the humans' play suggested wins with rock were preferred to wins with paper or scissors.

Given that people were sensitive to the payoff information the next question was whether, as in West & Lebiere (2001), the game theory solution predicted the move probabilities for the human players. Figure 1 displays the probabilities for playing paper, rock and scissors for the human subjects, for each of the opponents they faced, with 95% confidence intervals. Figure 1 also displays the predicted probabilities from the game theory solution. As can be seen, the game theory solution was significantly different from the human probabilities.

Table 1: The game results of the human versus the neural network models.

Network Model	Play ratios (P,R,S)	Mean points difference	Expected points difference	Strategy Index
Lag 1	28.8%, 39.6%, 31.5%	16.5	4.86	11.64
Lag 2	32.3%, 38.1%, 29.5%	5.7	11.38	-5.68
Rock=2 lag 1	33.5%, 39.7%, 26.8%	25.6	5.76	19.84

Table 2: Regression analysis on the performance of human subjects against the three network models.

Network Model	Regression coefficient	R Squared	Confidence Intervals Lower 95%	Upper 95%
Lag 1	0.0886	0.0208	0.0835	0.0938
Lag 2	0.0212	-0.0163	0.0160	0.0264
Rock=2 lag 1	0.0539	0.0251	0.0488	0.0590

We also examined whether the human results could be explained by using the game theory player model without the optimal game theory solution. That is, did the human subjects win by using move probabilities that exploited non-optimal move probabilities produced by the network models? Table 1 shows the average difference in score along with what the difference in score would be if it were determined solely by the overall move probabilities of the two players. The strategy index number is the difference between these two. A score of zero on the strategy index would indicate that the score difference could be accounted for entirely by move probabilities. In all cases the strategy index was significantly different from zero ($P<0.05$, determined by confidence intervals). Given the move probabilities, against the two lag 1 models the humans played significantly better than expected, while against the lag 2 model they played significantly worse than expected. This can be interpreted as the humans being better at exploiting sequential dependencies than the lag 1 models, but not as not as good as the lag 2 model. This agrees with the results of West & Lebiere (2001) and suggests that humans were able to narrowly beat the lag 2 because of their knowledge of the payoffs.

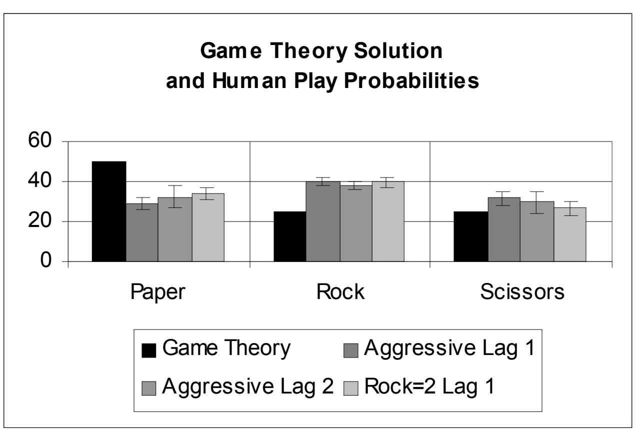

Figure 1: Game Theory Solution and Human Play Probabilities.

Modeling the Human results

Our next step was to construct a neural network model of how the humans played. For the model we assumed that people detect sequential dependencies in a way similar to a lag2 network. Although the results of this paper and West & Lebiere (2001) show that in games against the lag 2 network, humans seem to be slightly worse at detecting sequential dependencies, we again assumed this was due to humans finding the lag 2 less fun to play against because it is a stronger opponent. In both studies the advantage for the lag 2 network was relatively small. Additionally, West & Lebiere (2001) found that they could account quite well for the results of the other games by modeling humans as aggressive lag 2 networks. To account for the findings in this study we modeled people as aggressive lag 2 networks with the ability to adjust their rewards and punishments so as to best take advantage of the payoffs in the game.

To get an idea of how people could be adjusting the rewards and punishments we obtained self-reports on the strategies used by several of our more successful subjects. These reports generalized to favoring rock wins to paper wins, and paper wins to scissors wins. That is, they were focused first on getting rock wins and second on blocking the opponent from getting rock wins. With this in mind we ran a genetic algorithm to find a system of rewards and punishments for the neural network model that would match the human point difference results. The result was the following: rock wins = 3, paper wins = 2 scissors wins = 0; rock tie = -1, paper tie = -1, scissors tie = 0; and -3 for all losses. Note that not rewarding scissors wins makes sense as winning frequently with scissors would be associated with the opponent playing paper more often and that would block rock.

The model was played against each of the three networks that the humans faced. Each simulation consisted of 1000 games of 300 trials each. For each game, the net points difference between the models, the probabilities by which the human model selected each of the three play choices, and the strategy index for the human model was recorded. These three measures were used to determine how well the model fit the human data.

Results

The results showed that the model matched the point differences well (see Figure 2). The model also reproduced the human move probabilities against each opponent with a high degree of accuracy. The correlation between the model move probabilities and the human move probabilities was 0.964 ($p<0.0005$). Also, the model provided a good overall fit to the human strategy index data (see Figure 2). Although the model did not match the data as well when the opponent was the lag 2 network, this is actually consistent with our position that humans do not detect sequential dependencies as well against the lag 2 network due to confounding factors, which would not apply to the network model of human behavior. Also, the fact that the model was able to match two sets of results that it was not explicitly designed to match (the move probabilities and the strategy index values) suggests that the results were reasonably robust.

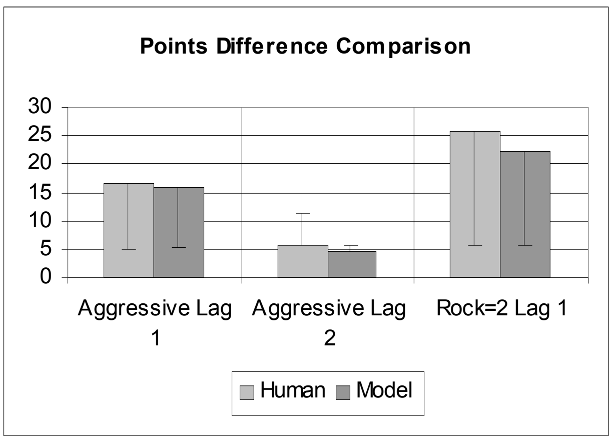

Figure 2: Points Difference Comparison. The shaded bars indicate the mean final points difference. The heads of the T-bars represent the expected final points difference; and, the length of the T-bars indicate the magnitude of the strategy index values.

Conclusions

The results of this study replicate the West & Lebiere (2001) findings that the commonly used probabilistic game theory model (as defined in this paper) cannot account for the game results when humans play against agents programmed to play by exploiting sequential dependencies. We also demonstrated that when the game payoffs are not all equal, the game theory solution does not predict the aggregate move probabilities. We further demonstrated, using the actual move probabilities, that the results could not be accounted for by the game theory player model. That is, the actual move probabilities did not predict the final points differences. These results show that the game theory player model, with or without the game theory solution, is fundamentally different from the way people process information in this type of situation.

In terms of modeling, we replicated the West & Lebiere (2001) result that this type of human game playing can be accurately modeled using simple lag 2 networks. Furthermore, we extended the original model by showing that people are sensitive to different game payoffs and that this can be modeled by adjusting the rewards and punishments associated with different play outcomes. These results are also consistent with a number of ACT-R studies showing that people play a variety of games using the lag 2 strategy (PRS: Lebiere & West, 1999: non-zero-sum games: Lebiere, Wallach, & West, 2000; baseball: Lebiere, Gray, Salvucci, & West, 2003). The ACT-R model works by using the ACT-R declarative memory system as a neural network for detecting sequential dependencies, and produces results similar to the simple networks we used (Lebiere & West, 1999). An ACT-R model equivalent to the one in this paper could be created by “popping” the “chunks” representing sequential dependency patterns a different number of times for different outcomes. Likewise, a genetic algorithm could be used to fit the model. However, this approach would sidestep the next important issue, which is modeling how humans adjust their reward structure in response to the game payoffs.

References

Estes, W. K. (1972). Research and theory on the learning of probabilities. *Journal of the American Statistical Association*, 67, 81-102.

Fudenburg, D., & Levine, D. K. (1998). *The theory of Learning in Games*. Cambridge, MA: The MIT Press.

Gazzaniga, M. S. (1998). The split brain revisited. Scientific American, July, 50-55.

Lebiere, C., Gray, R., Salvucci, D., & West, R. L. (2003). Choice and Learning under Uncertainty: A Case Study in Baseball Batting. *Proceedings of Cognitive Science*.

Lebiere, C., Wallach, D., & West, R. L. (2000). A Memory-based Account of the Prisoner’s Dilemma and Other 2x2 Games. *The proceedings of the Third International Conference on Cognitive Modeling.*

Lebiere, C., & West, R. L. (1999). Using ACT-R to model the dynamic properties of simple games. *Proceedings of Cognitive Science*.

Pool, R. (1995). *Putting game theory to the test*. Science, 267, 1591-1593.

Restle, F. (1966). Run structure and probability learning: Disproof of Restle’s model. *Journal of Experimental Psychology*, 72, 382-389.

Ritter, F. E., & Wallach, D. P. (1998). Models of two-person games in ACT-R and Soar. *Proceedings of the Second European Conference on Cognitive Modelling*. 202-203. Thrumpton: Nottingham University Press.

Rose, R. M., & Vitz, P. C. (1966). The role of runs of events in probability learning. *Journal of Experimental Psychology*, 72, 751-760.

Samuelson, L. (1997). *Evolutionary games and equilibrium selection*. Cambridge, Mass: MIT Press.

Poundstone, W. (1992). *Prisoner’s dilemma*. NY: Doubleday.

Vitz, P. C., & Todd, T. C. (1967). A model of learning for simple repeating binary patterns. *Journal of Experimental Psychology*, 75, 108-117.

VonNeumann, J., & Morgenstern, O. (1944) *Theory of games and economic behaviour*. Princton, N. J.: Princton University Press.
Ward, L. M. (1973). Use of markov-encoded sequential information in numerical signal detection. *Perception and Psychophysics*, 14, 337-342.
Ward, L. M., Livingston, J. W., & Li, J. (1988). On probabilistic categorization: The markovian observer. *Perception and Psychophysics*, 43, 125-136.
Wagenaar, W. A. (1972). Generation of random sequences by human subjects: A critical survey of the literature. *Psychological Bulletin*, 77, 65-72.
West, R. L. (1998). Zero Sum Games as Distributed Cognitive Systems. *Proceedings of the Complex Games Workshop*. Tsukuba, Japan: Electrotechnical Laboratory Machine Inference Group.
West, R. L., & Lebiere, C. (2001). Simple games as dynamic, coupled systems: Randomness and other emergent properties. *Cognitive Systems Research*, 1(4), 221-239.