Accounting for the integration of descriptive and experiential information in a repeated prisoner's dilemma using an instance-based learning model

Noam Ben-Asher*, Varun Dutt*, & Cleotilde Gonzalez*

*Dynamic Decision Making Laboratory, Carnegie Mellon University
4609 Winthrop Street, Pittsburgh, PA 15213
noamba@cmu.edu, coty@edu.cmu.edu

*School of Computing and Electrical Engineering, School of Humanities and Social Sciences, Indian Institute of Technology, Mandi, India.
varun@iitmandi.ac.in

Keywords:
Cognitive modeling, Instance-based learning, repeated prisoners dilemma, descriptive information.

ABSTRACT: The way information is presented, using description or experience, can influence the decision making process. However, little is currently known on how descriptive information is accounted for in subsequent experiential learning. In this paper, we use a computational model based upon Instance-based Learning Theory (IBLT) and use it to study hypotheses on how participants may integrate descriptive and experiential information presented in a repeated prisoner’s dilemma. Two players, each simulated by an Instanced Learning (IBL) model, play against each other in a repeated prisoner’s dilemma. They are provided with a descriptive payoff matrix as well as the experiential information in the form of each other’s payoffs after making repeated decisions. Our results demonstrate that the descriptive payoff matrix information is incorporated into decisions using two mechanisms. The first is expectations derived from the description. A second mechanism, highlights the immediate and worst possible outcomes in the description, giving them more attention compared to the other outcomes. We highlight the significance of our results for decision making in social dilemmas.

1. Introduction

Often in real-life, we make decisions while accounting for information coming from different sources. In situations involving conflict, for example, one can decide whether or not to cooperate with a partner, based on descriptive information on the benefits of each of the two people involved in the transaction (i.e., a payoff matrix). However, one can also use past experience of interacting with that or other partners to find out whether to cooperate or not in a particular situation.

Studies of individual decision making distinguish between these two sources of information (descriptive and experiential), and demonstrate how decisions from experience and description differ (Barron & Erev, 2003; Hertwig, Barron, Weber, & Erev, 2004). Recently, a similar distinction has been made in social dilemmas and games of cooperation such as the Prisoner's Dilemma (PD, see Martin, Gonzalez, Juvina, & Lebiere, 2012 for a discussion of the Experience-Description gap in social dilemmas).

However, current studies often make a sharp contrast between descriptive and experiential information ignoring the fact that in real-life, both types of information are often available at the same time. As a result, much less is known about how decisions are made when both types of information are present (Newell & Rakow, 2007; Jessup, Bishara, & Busemeyer, 2008; Lejarraga & Gonzalez, 2011), particularly in situations of conflict and cooperation (i.e., social dilemmas) (Gonzalez & Martin, 2011; Martin et al., 2012).

Gonzalez and colleagues (Gonzalez, Ben-Asher, Martin, & Dutt, 2012) have attempted to address this gap in the literature by creating a cognitive model (called “IBL-PD”) of experiential decision making in the Prisoner's Dilemma that is based on the Instance-Based Learning Theory (IBLT) (Gonzalez, Lerch, & Lebiere, 2003). This model was used to test different hypotheses regarding how a player accounts for the partner's information (both descriptive and experiential) in a repeated PD game. They conclude that a dynamic regard of the partner's information based on surprise (gap between expected outcomes and actual outcomes) best explains the effects of experiential information. The IBL-PD model was extended to account for descriptive information by explicitly directing attention to the description of the worse outcomes.
Although Gonzalez et al. (2012) present significant contributions to the current knowledge of decisions from experience and description in social dilemmas, there are still many unanswered questions regarding how descriptive information is integrated and merged with experiential information to make decisions. It is possible that initially in the absence of prior experiences, descriptive information is mainly used to form expectations regarding beneficial strategies that a player should adopt. Such expectations become irrelevant with the availability of experiences and the influence of the initial expectations decrease as the number of experiences increase. It is also possible that in addition to the long-term expectations, the descriptive information is explicitly encoded in memory. However, this alternative raises a challenging question: do people encode all pieces of information equally or do they pay more attention to the salient outcomes? For example, do we highlight the description of the worse possible outcome compared to all other outcomes, or does the best possible outcome gets more attention?

The current paper addresses alternative ways of integrating experiential and descriptive information in the repeated PD using behavioral data collected in past research (reported in Martin et al., 2012) and an IBL-PD model created in Gonzalez et al. (2012). We start by providing background information on a repeated PD game that was used to collect behavioral data, and present an IBL-PD model that accounts for experiential information in the game. Then, we position four competing hypotheses regarding the integration of descriptive and experiential information within the mechanisms proposed by this learning model.

2. Effects of Descriptive and Experiential Information in the Prisoner's Dilemma

Game theory is concerned with the decisions of players who are aware that their actions affect each other. One objective for a player in a 2x2 game (between two players, each of whom has two available choices), is to maximize personal economic benefit. The Prisoner's Dilemma (PD) (Axelrod, 1980; Rapoport & Chammah, 1965) has been widely used to investigate such conflict situations. In the PD, each of the two players decides without communication, whether to cooperate (C) or defect (D). Following, the outcome of each player is determined by combining the decisions of the two players. In this game, defection leads to higher expected outcome regardless of the other player decision (see payoffs in Figure 1). For a one-shot PD game, the standard finding is that a larger proportion of defection over cooperation, when aggregated over many participants (Rapoport & Chammah, 1965).

![Figure 1. Prisoner's Dilemma payoff matrix, with Option D denoting defection and Option C denoting cooperation. The cells show a pair of outcomes (x, y) where x is the payoff to Player 1 and y is the payoff to Player 2.](image)

In a repeated PD, two players are asked to repeatedly decide without communication whether to cooperate or defect across multiple rounds. Players face a tradeoff between the immediate gain of defection and the potential long-term benefits of mutual cooperation. The latter can only be achieved if both players cooperate. The standard finding is that the proportion of defection initially increases and then starts to decrease over time as mutual cooperation increases (Rapoport & Chammah, 1965; Martin et al., 2012). This finding emphasizes the importance of repeated experience in the PD, and demonstrates how the benefit of long-term mutual gain of cooperation eventually overcomes the attractiveness of immediate individual gain from defection (Baker & Rachlin, 2002).

In a repeated PD, a player has information from two different sources. The payoff matrix which provides a description of the possible options a player has to choose from and their corresponding outcomes. In repeated PD, the two players see the descriptive information before making the first decision and throughout the game. On the other hand, experiential information is available to the players, but only after making a decision and observing the opponent's decision and their mutual outcomes. Based on previous decisions the opponent made and what were the outcomes of these decisions, a player can conclude regarding the tendency of the opponent to prefer one option over the other. Furthermore, each time a decision is made and the player receives feedback, the amount of available experiential information increases. This highlights an important dependency between the two types of information. The descriptive information is always presented as part of the environment and does not change over time. In contrast, the player is accumulating a growing amount of experiential information in memory. Thus, the amount of experiential information can influence how a player attends to the descriptive information.

Martin et al. (2012) examined the effects of information on cooperation in repeated PD. Their findings indicate that when players have both descriptive and experiential information, the mutual cooperation is higher compared to a situation where only experiential information is available to the players. To highlight the differences in behavior related
to the availability of descriptive information, we summarize here their methods and some of their findings.

One-hundred and twenty participants were randomly paired and assigned to play 200 trials of repeated PD in one of two between-subjects information conditions. In each round of the Experiential condition, a player saw the outcome from her decision and also saw the action and outcome of the other player. In the Descriptive condition, the player had the same information as in the Experiential condition, in addition to the complete payoff matrix (as in Figure 1), was presented from the outset and throughout the repeated interaction.

As seen in Figure 2, the proportion of cooperation in the two information conditions follows a similar pattern: initially there is a decrease in the tendency to cooperate and as experience in the game increase there is gradual increase in the proportion of cooperation. However, the proportion of cooperation in the Descriptive condition is consistently and significantly higher than the proportion of cooperation in the Experiential condition, \(t(199)=30.55, p<0.001\).

![Figure 2. The proportion of cooperation over time for human participants in the Experimental and Descriptive conditions.](image)

The source of this gap is the availability of descriptive information in addition to the experiences gathered during the interaction between the two players. In this paper, we investigate possible ways to account for this systematic gap between the two information conditions that expand those reported in Gonzalez et al. (2012). We examine different ways of incorporating descriptive information in an experiential learning model by using the IBL-PD model, reported in Gonzalez et al. (2012). This model is explained next for completeness. Finally, we discuss the results of the comparison and their implications.

### 3. IBL-PD Model

According the IBLT, choices are made based on past experience and according to what has led to the best outcomes in similar situations in the past. The IBL model relies on a set of mechanisms adopted from the well known cognitive architecture, ACT-R (Anderson & Lebiere, 1998, 2003).

From the players' perspective, repeated PD can be seen as a repeated binary-choice between cooperation and defection. Hence, two models with identical mechanisms and parameters can represent two players, playing PD repeatedly. In any given trial, each model makes a decision independently, and the outcomes are determined as a consequence of the two decisions together, as for human players.

The key representation of cognitive information in the IBL model is an instance. An instance represents the connection between a decision and its outcome in the following form [Decision, Player’s Outcome, Opponent’s Outcome]. Thus, for each player in the repeated PD game there are four possible instances: [C, -10, 10], [C, 1, 1], [D, 10, -10], [D, -1, -1], representing two possible decisions (C or D) and four possible outcomes. According to IBLT, in the absence of prior information (i.e., description) these instances are created only after an outcome is experienced, and are reinforced whenever they are re-experienced. After an instance is created, it can be retrieved and used to make decisions, according to the following decision making process.

In the IBL model, the selected option in a trial is the one with the highest "blended" value (Lebiere, 1999; Gonzalez et al., 2003; Gonzalez & Dutt 2011), resulting from averaging all instances belonging to a decision option. In the IBL-PD model, the blended value of option \(j\) is defined as:

\[
V_j = \sum_{i=1}^{n} p_{ij}(x_{ij} + w_{oij})
\]

(1)

Where \(x_{ij}\) is the outcome of the player stored in an instance \(i\) for option \(j\), and \(o_{ij}\) is the opponent's outcome corresponding to the player's selection of option \(j\) and observing outcome \(x_{ij}\). The \(p_{ij}\) represent the probability of retrieving instance \(i\) from memory for blending purposes. The \(n\) is the number of different outcomes observed by selecting option \(j\), up to the last trial. The "weight" \((w)\) represents the extent to which a player is willing to consider the corresponding outcome of the opponent \((o_{ij})\) when selecting option \(j\) at time \(t\).

The best account of the repeated proportion of cooperation in the human data (see Gonzalez et al., 2012 for a detailed description) is an IBL-PD model.
that assumes that the weight a player gives to the opponent's outcome is a function of the gap between the player's expected outcome and the outcome actually received (i.e., surprise). Thus, the value of $w$ changes over time by as a function of surprise:

$$w_t = 1 - \text{Surprise}_t$$  

(2)

The following definition of $\text{Surprise}_t$ as per Erev, Ert, and Roth (2010) and Gonzalez, Dutt, and Lejarraga, (2011) was adopted by Gonzalez et al. (2012) to account for the information provided by experiencing the opponent's outcome in repeated PD:

$$\text{Surprise}_t = \frac{\text{Gap}_t}{[\text{Mean}(\text{Gap}_{t}) + \text{Gap}_t]}$$  

(3)

Where the Gap at time $t$ after choosing option $j$ and observing the outcomes stored in instance $i$ is defined by:

$$\text{Gap}_t = |V_j - (x_{ij} + a_{ij})|$$  

(4)

The Gap is the absolute difference between the blended value of option $j$, representing an expected weighted outcome, and the outcomes that were actually received by selecting option $j$.

The mean Gap is defined assuming a horizon of 200 trials of the repeated PD as follows:

$$\text{Mean}(\text{Gap}_{t}) = \text{Mean}(\text{Gap}_{t-1}) \left(1 - \frac{1}{200}\right) + \text{Gap}(t)\left(\frac{1}{200}\right)$$  

(5)

Equation 3, indicates that the value of $\text{Surprise}_t$ and as a result, the value of $w$, are between 0 and 1. The higher the value of $w$ is the more weight is given to the opponent's outcome.

The probability of retrieving an instance $i$ from memory is a function of its activation ($A_i$) relative to the activation of all other instances that correspond to option $j$. In each trial $t$, the retrieval probability is defined as:

$$p_{ij} = \frac{A_i}{\sum_i A_i}$$  

(6)

Where $\tau$ is random noise defined as $\sqrt{2}\sigma$, and $\sigma$ is a free noise parameter that accounts for the imprecision of recalling instances from memory for blending (adapted from the ACT-R architecture; Anderson & Lebiere, 1998, 2003). Thus, a high $\sigma$ value implies a greater variability in retrieving instances from memory.

The activation of each instance in memory depends upon the Activation mechanism originally proposed in the ACT-R architecture. The IBL-PD model uses a simplified version of that mechanism which relies on recency and frequency of use of instances (Lejarraga, Dutt, & Gonzalez, 2012). In each trial $t$, activation of instance $i$ is:

$$A_i = \ln\left(\sum_{t \in \{1, \ldots, t-1\}}(t - t_i)^{-d} + \sigma \cdot \ln \left(\frac{1-Y_i}{Y_i}\right)\right)$$  

(7)

Where decay $d$ is a free parameter and $t_i$ (i.e., a timestamp) refers to the previous trial in which the outcome was observed. Therefore, the activation of an instance containing an outcome increases with the frequency of observing that outcome as well as with the recency of observing that outcome. The $d$ parameter accounts for the rate of forgetting information: A high value of the $d$ parameter leads to a faster decay of an instance in memory. The $Y_i$ term is a random draw from a uniform distribution $U(0,1)$, and the $\sigma \cdot \ln \left(\frac{1-Y_i}{Y_i}\right)$ term represents the Gaussian noise for capturing the participant-to-participant variability in activation.

### 3.2. The role of the first trial in the IBL-PD model

When learning only from experience, instances are created by making a decision and observing its outcome. However, when making the first decision in the repeated PD game, players have only descriptive information. This information can influence their first decision by generating some expectations regarding the opponents’ action in the game. The IBL-PD model assumes that there are two pre-populated instances in memory corresponding to the two options the decision maker has. The values of the pre-populated instances represent the decision maker’s initial expectations before any experience is acquired in the task (Gonzalez et al., 2011). When modeling learning from experience in a binary-choice, Lejarraga et al. (2012) used the pre-populated instances to trigger initial exploration of two decision options. In a similar manner, Gonzalez et al. (2011) adjusted the values of the pre-populated instances to influence the decision of an IBL model in the first trial of a Market Entry Game.

Overtime, as the values stored in the pre-populated instances are not observed in the PD game and with the arrival of new experience, the activation of the pre-populated instances generally decay. Their decay leads to a decrease in the retrieval probability, which in turn leads to a decrease in the contribution of these values to the calculation of the blended value. However, the pre-populated instances may have a long-term memory trace that depends on the combination of their value and activation. In the IBL-PD, we use the pre-populated instances to represent the expectations a player has from perusing the descriptive information that the payoff matrix provides. We derive the values in pre-populated instances by calibrating our model to the proportion of cooperation across trials.
4. Integrating Descriptive and Experiential Information in the IBL-PD Model

Several studies demonstrated that displaying the complete payoff matrix to repeated PD players increased the overall level of cooperation compared to a condition where the payoffs were learned through repeated experiences only (Rapoport, Guyer, & Gordon, 1976; Martin et al., 2012). However, the mechanism through which the descriptive information alters the players’ behavior is not clear. It is possible that when the descriptive information is observed for the first time, it creates some expectations regarding beneficial strategies that a player should adopt and the possible strategies that the opponent might use. More generally, it allows the players to reason through the underlying dependencies between their decisions and outcomes, including the forgone payoffs for each of the possible decisions (Gonzalez & Martin, 2011).

In light of the assumptions above, one can ask what is the impact of the descriptive information beyond the first decision? Especially, how do the expectations integrate with the new information that the feedback provides after each trial of repeated PD? Studies suggest that experiencing feedback following a decision alters the interpretation of the description in repeated choice tasks. For example, Barron, Leider, and Stack (2008) found that positive experiences in the beginning of the interaction increased risk taking even in the presence of a descriptive warning. Similarly, Jessup et al. (2008) showed that feedback influenced decisions from description in binary choice task, leading to underestimation of the probability to experience a rare outcome. Finally, the findings in Lejarraga and Gonzalez (2011) indicate that decision makers rely heavily on experience even when they are provided with accurate descriptive information. This behavior of human decision makers was accurately replicated by two different cognitive models (IBL and reinforcement learning) that had no representation of the descriptive information.

Four distinct hypotheses may be drawn regarding how to account for descriptive information in IBL-PD and what information is obtained from the payoff matrix:

1. The first hypothesis implies that the descriptive information is translated into prior expectations that are used to make a decision before any experiential information exists (Gonzalez & Martin, 2011). This hypothesis follows the idea that information from description is quickly neglected in the presence of feedback (Lejarraga & Gonzalez, 2011). According to this hypothesis, the descriptive information is represented in the IBL-PD model in the values of the pre-populated instances only. These instances represent the expectations that a player has before any information regarding the opponent behavior is available. Other instances that represent the actual combination of decision and outcome are only created based upon experience.

2. A competing hypothesis can be drawn from the work of Jessup et al. (2008), suggesting that the information from description and from experience are combined in the decision process. Therefore, in a corresponding IBL-PD model, expectations are derived from descriptive information and stored in the pre-populated instances and at the same time the descriptive information generates representations of the immediate outcomes by creating instances in memory that correspond to all of the possible decisions and outcomes based on the payoff matrix. Thus, when IBL-PD makes the first decision, the activation of the pre-populated instances and the activation of the instances generated from the matrix are equal, and all forms of information are considered equally in the calculation of the blended value. This hypothesis also suggests that all the descriptive information included in the matrix receives the same attention before making the first decision.

3. Building upon Prospect Theory and loss aversion, (Kahneman & Tversky, 1979), it is possible to assume that the option that yields the highest gains (i.e., +10) or the highest losses (i.e., -10) in the descriptive information are weighted differently in the decision process. Thus, in contrast to the idea that all the descriptive information is activated equally when making the first decision, it is possible that only the best or worst outcomes are highlighted in memory and form the short-term expectations. This assumption leads to additional two hypotheses: first, that there are expectation stored in the pre-populated instances in the IBL-PD and only the instance that represents the best outcome is created before making the first decision.

4. Alternatively, and as suggested by Prospect Theory, losses should receive higher weight than gains. An IBL-PD model that follows this hypothesis will have expectations stored in the pre-populated in stances and only the instance representing the worst outcome is created before making the first decision (Gonzalez et al., 2012).

We test four IBL-PD models, each correspond to a competing hypothesis stated above. For all models the values of the free parameters $\sigma$ (noise) and $d$ (decay) were set at 1.5 and 5, respectively (obtained from a different data set in Lejarraga et al., 2012). We used a Genetic Algorithm (Swisher, Hyden, & Jacobson, 2000) to calibrate the values of the pre-populated instance for each hypothesis separately. The goal of the calibration was to minimize the mean squared distance (MSD) between the model's proportions of cooperation
and the corresponding human’s proportions of cooperation in each trial.

Performance was determined by computing the MSD and correlation over 200 trials between the cooperation rate predictions from the model with 100 simulated pairs and that from the human data. The model with the lowest MSD and the highest correlation will provide evidence regarding the way descriptive information can be accounted for in a repeated PD.

5. Results

The results illustrated in Figure 3 present the cooperation rate of humans and the IBL-PD model under the assumptions of the first hypothesis. In the model that corresponds to the first hypothesis, we only calibrated the values stored in two pre-populated instances, one associated with defection (D) and the other with cooperation (C). The MSD between human data and model prediction was minimized at 0.0109 when pre-populated instance D = -0.14 and pre-populated instance D = -0.82. When representing the descriptive information only in the two pre-populated instances, the model yields a relatively low MSD. However, the low and insignificant correlation ($r=0.08$, $p=ns$) between human data and model predictions indicate that the model did not capture the dynamics within cooperation rate over time.

When making the first decision at the first trial, the models had only the pre-populated instances, and the first decision was made solely based on their values. Thus, unlike humans, the model always defected in the first trial as defection seemed more attractive than cooperation based on the values of the pre-populated instances. After making the first decision and with the increased availability of feedback, the cooperation rate increases relatively quick and stabilize, unlike human behavior.

For the second hypothesis, the model had, in addition to calibrated pre-populated instances (D = 2.09, C = 30.64), 4 additional instances that represent all the possible decisions and outcomes based on the payoff matrix. These instances were created at the start based on the descriptive information before the model had any experience. When experiences were available to the model, the activation mechanism (see Equation 7) reinforced these existing instances that corresponded to the combination of a decision and an outcome. As seen in Figure 4, combining expectations (through pre-populated instances) and explicit information regarding the immediate outcomes changed the model’s behavior when making the first decision and over time.

When calculating the blended value (Equation 1) for the first time, the model equally considers the expectations and the possible immediate outcomes because the activation of all the instances is relatively similar and it differs only due to the Gaussian noise as per Equation 7. This, together with the initial high expectations from cooperation, replicates well humans decision in the first trial. However, with the availability of feedback and as the activation of the instances generated from the matrix increase while the activation of the pre-populated ones decrease, we find a sharp decrease in the proportion of cooperation. Only after the model learns the disutility of mutual defection, cooperation emerges. The prediction of this model has a slightly higher MSD (0.0198) compared to the previous model. However, the main benefit from using the descriptive information to generate all the instances from the outset is the high and significant correlation between human data and model predictions, $r=0.73$, $p<0.001$.

Next, we tested the third hypothesis that descriptive information generates expectations stored in the pre-populated instances and, in addition, an instance in memory from the start that represents the immediate highest possible gain. This highlights the outcome of +10, which can be achieved when a player defects and the opponent cooperates. According to this model, all the other instances are created from experience. As before, we calibrated the values of the pre-populated instances representing expectation from defection (D = -22.68) and cooperation (C = 30.32).

The model corresponding to the hypothesis that the best outcome is highlighted yield a relatively similar MSD (0.0201) and correlation ($r=0.76$, $p<0.001$) as the model that had four instances representing all the descriptive information from the first trial. However,
the two models perform differently when making the first decisions.

Figure 4. Proportion of cooperation in human data and model prediction. The model uses the descriptive information to derive both expectations and representations of all immediate outcomes.

As seen in Figure 5, the current model always cooperates when making the first decision, even though the representation of the descriptive information highlights the high and positive outcome from defection (+10). It seems that the first decision is mostly influenced by low value of the pre-populated instance that represents defection. However, as activation of the pre-populated instances decay over time, the proportion of cooperation decreases. Similar to the previous model, with the experience of the outcomes of mutual defection, the model gradually learns to cooperate and the proportion of cooperation picks up.

Figure 5. Proportion of cooperation in human data and model prediction. The model uses the descriptive information to derive expectations and highlights the representation of the best immediate outcomes.

The last hypothesis suggests that the descriptive information generates expectation for both options and highlights the worst immediate outcome. This outcome of -10, is achieved when a player cooperates and the opponent defects. Following the same procedure, we created the instance representing the worst outcome before making the first decision; and, also calibrated the pre-populated instances representing defection and cooperation.

As seen in Figure 6, this model predicts very well human behavior in the first trial and over 200 trials (this result is also reported in Gonzalez et al., 2012). The MSD (0.0026) obtained by this model with the repopulated values of $D = -2.46$ and $C = 13.86$ is much smaller compared to the previous modes, and at the same time the correlation between the model’s prediction and human data is significant and high, $r=0.79, p<0.001$. The results suggest that it is possible to account for descriptive information in an experiential learning model using pre-populated instance that represent prior expectations and an explicit representation of the worst outcomes.

Figure 6. Proportion of cooperation in human data and model prediction. The model uses the descriptive information to derive expectations and highlights the representation of the worst immediate outcomes.

6. Discussion and Conclusion

In this paper, we investigated how to combine information from description and from experience in IBL model of a repeated PD player. Four different hypotheses were evaluated, and the most promising one suggest that descriptive information is incorporated with experience using two distinct mechanisms. The first mechanism provides expectations from different strategies and the second highlights the worst immediate outcome.
The expectations can be seen as beliefs regarding actions of the opponent and the strategy the player should use. These expectations are made without any experience and as such their relevance decay as experience provide more accurate information regarding the environment, including the opponent’s actions. However, the expectations are not ignored immediately at the precedence of experiential information and they seem to have a long-term memory decay (d).

The second mechanism follows the idea that not all parts of the descriptive information are incorporated equally. While the pre-populated instances provide expectations derived from the description, the worst immediate outcome in the description receive special attention. The IBL model attends to the worst possible outcome, even before the first decision is made. This notion is in agreement with Prospect Theory and loss aversion (Kahneman & Tversky, 1979), as losses are weighted more than gains.

This work opens a way to use IBL modeling in environments that incorporate information from description and from experience. It allows the IBL model to incorporate the descriptive information into the decision making process without having to program specific behavioral rules or state predefined strategies. This will facilitate IBL modeling in the future, and can help to understand the source of discrepancy between decisions from description and from experience.

Our current implementation, however, still has limitations and future work should evaluate them carefully. Currently it is not clear how the descriptive information is transformed to the exact values stored in the pre-populated instances. In this study, these values were obtained from human data; however, it is possible that there is a way to formulate this transformation. Furthermore, while the worst outcome is highlighted, it is possible that other possible outcomes are highlighted, but not to the same extent. Further research is required to determine the way attention is distributed across multiple possible outcomes in the payoff matrix: Whether participants pay attention to the most negative outcome from the start (as our results suggest); or if they pay attention to all outcomes from the start but in unequal amounts.

7. Acknowledgements

This research was supported by a Defense Threat Reduction Agency (DTRA) grant to Cleotilde Gonzalez and Christian Lebiere (HDTRA1-09-1-0053); and by the U.S. Army Research Laboratory under Cooperative Agreement No. W911NF-09-2-0053. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on.

8. References


Author Biographies

NOAM BEN-ASHER

Noam Ben-Asher is a Postdoctoral Fellow at the Dynamic Decision Making Laboratory at Carnegie Mellon University. His current research interests focus on dynamic decision making, and modeling of human behavior. He is particularly interested in the tradeoffs and interactions between risk taking, security, fairness, and performance in conflict situations like cyber warfare and social dilemmas.

VARUN DUTT

Varun Dutt is Assistant Professor, School of Computing and Electrical Engineering, School of Humanities and Social Sciences, Indian Institute of Technology (IIT), Mandi, India. His research is at the intersection of computer science, economics, and decision making, with a special focus on how computational and experimental models could be used for taking decisions on managerial, environmental, and socio-economic issues. He has so far published 21 journal articles, including two in Psychological Review, three in Journal of Behavioral Decision Making, and one in Climatic Change. Dr. Dutt also serves as the Knowledge Editor of the Financial Chronicle, a financial daily published from India.

CLEOTILDE GONZALEZ

Cleotilde Gonzalez is Associate Research Professor in the department of Social and Decision Sciences at Carnegie Mellon University. She is affiliated faculty with the Human-Computer Interaction Institute, the Center for Cognitive Brain Imaging and the Center for Neural Basis of Cognition, all at Carnegie Mellon University. She is also the founding director of the Dynamic Decision Making Laboratory (http://www.cmu.edu/ddmlab). Her interdisciplinary research program focuses on the study of human decision making in dynamic and complex environments. She has developed a cognitive theory of decision making in dynamic tasks, Instance-Based Learning Theory (IBLT), to explain and predict experiential-based choice behavior.