

Special Issue: New Theoretical Perspectives on Conflict and Negotiation

Scaling up Instance-Based Learning Theory to Account for Social Interactions

Cleotilde Gonzalez and Jolie M. Martin

Dynamic Decision Making Laboratory, Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA, U.S.A.

Keywords

social information, conflict, instance-based learning theory, Prisoner's Dilemma.

Correspondence

Cleotilde Gonzalez, Dynamic Decision Making Laboratory, Carnegie Mellon University, Pittsburgh, 15213 PA, U.S.A.; e-mail: coty@cmu.edu.

Abstract

Traditional economic theory often describes real-world social dilemmas as abstract games where an individual's goal is to maximize economic benefit by cooperating or competing with others. Despite extensive empirical work, descriptive models of human behavior in social dilemmas are lacking in both cognitive realism and predictive power. This article addresses a central challenge arising from the success of modeling individuals making decisions from experience: our ability to scale these models up to explain social interactions. We propose that models based on the instance-based learning theory (IBLT) will help us to understand how conflictual social interactions are influenced by prior experiences of involved individuals and by information available to them during the course of interaction. We present mechanisms by which IBLT might capture the effects of social interaction with different levels of information without assuming predefined interaction strategies, but rather by assuming learning from experience.

Introduction

Conflicts are a natural part of social interactions. Individuals may perceive contradictions between their own interests and those of others. Psychologically, decision makers strive to enhance their own self-interests while maintaining a sense of fairness and

This research was supported in part by a grant from the Defense Threat Reduction Agency (DTRA) grant number: HDTRA1-09-1-0053 to Cleotilde Gonzalez and Christian Lebiere. We thank Hau-yu Wong for editing this manuscript and Christian Lebiere and Ion Juvina for discussions that contributed to ideas in this article.

obligation to others at the same time. Like many everyday decisions, decisions in situations of conflict often require a delicate balance between furthering one's own interests and those of others. Complicating matters further, these interests must be estimated from personal experience and limited information about others' motivations. The study of decisions from experience (DFE) is an important new shift of attention in decision sciences away from the traditional study of decisions from description (Hertwig, Barron, Weber, & Erev, 2004). DFE represents a natural way of studying how humans adapt and learn to make decisions in the absence of explicitly stated payoffs and probabilities of obtaining those payoffs. And while classical behavioral decision theory focuses on violations of rationality assumptions, DFE helps to understand human behavior in natural environments where one cannot rely on the rationality assumption because of social complexity (Erev & Haruvy, in press).

Our research addresses a central challenge arising from the success of models of individual DFE: scaling these models up to represent the response of decision makers to social information. We offer a model of individual DFE that can be expanded to account for the repeated effects of social interactions. We propose that behavior in conflict is determined by the availability of relevant social information and by the combined effects of past experience in similar situations. The presence of information regarding others' interests and outcomes will drive attention toward particular beliefs, attitudes, and motivations, which influence conflict resolution. The model is based on the instance-based learning theory (IBLT; Gonzalez, Lerch, & Lebiere, 2003), which has proven to be a robust representation of DFE in individuals.

Our experimental approach to collecting human data against which the model can be compared is derived from game theory, which often describes real-world social dilemmas as abstract games where an individual's goal is to maximize economic benefit by cooperating or competing with others. For example, in the well-known Prisoner's Dilemma (PD; Axelrod, 1980; Rapoport & Chammah, 1965), each of two players chooses to "cooperate" or "defect" given information about the payoffs if one or the other action are chosen by each of the players. In the Iterated Prisoner's Dilemma (IPD) where players interact repeatedly, computing the best response to maximize total payoffs becomes more complex as additional strategic factors come into play, including the possibility for the other to punish defection or reciprocate cooperation (Rabin, 1993).

In contrast to a purely economic perspective, a behavioral perspective explains how conflict and cooperation are impacted by social dimensions such as interpersonal relationships, past encounters, identities, and emotions (Roberts, 1997; Schuster, 2000; Schuster & Perelberg, 2004); it is important to disentangle the influence of these social concerns above and beyond monetary outcomes for individuals (e.g., Colman, 2003). Drawing upon experience in the real world, people may implicitly care about an ongoing relationship and the image they project to others, despite being told to disregard these factors. That is, people make DFE in social interactions, and feedback may reinforce some social outcomes and behaviors over others. For example, deviations from the mutual defection equilibrium in the one-shot PD could result from inherent altruism, as well as from norms prescribing fairness and kindness. Just as economic

considerations become more complex in an IPD played for multiple rounds, so too do social considerations become more complex in repeated interaction.

One important social aspect of conflict relates to the information available to decision makers, which ranges from no information about interdependence with others to complete information about the actions of others, their influence on the other's outcomes, and the cultural identities of others. These informational characteristics influence reciprocity, fairness, trust, power, and other social concerns. Not surprisingly, the effects of these social factors on individual and collective behavior are difficult to capture in computational models that have oftentimes been designed to account primarily for individual decisions and outcomes. We organize the information that is expected to influence social interactions into a *Hierarchy of Social Information* (hereafter HSI). The HSI framework helps conceptualize different levels of social information, organize empirical finding from the literature that demonstrate the influence of social information, and propose systematic ways in which IBLT can be extended to capture these empirical effects. Ultimately, our objective is to develop a generic IBLT model that accurately represents the major influences of social information on cognition and emotion in conflict and is thus capable of predicting behavior in novel contexts.

In what follows, we first review models that explain how individuals make decisions from experience. We then describe IBLT, the current assumptions and mechanisms of the theory, and discuss ways that IBLT might be extended to represent different levels of social information in the HSI.

Models of DFE

Many models have explained individual behavior in DFE. Learning models often emphasize a weighted adjustment process by which the value of a previously observed outcome is combined or updated with a newly observed outcome (e.g., Bush & Mosteller, 1955; Erev, Glozman, & Hertwig, 2008; Hertwig, Barron, Weber, & Erev, 2006; March, 1996). Reinforcement models of learning are perhaps even more common in the literature. These models often assume that choices are reinforced based on immediate feedback (Erev & Roth, 1998; Roth & Erev, 1995). Authors of these models make cognitive or psychological attributions to different parameters of their models. Attributions such as attention and learning ability are common (Ho, Camerer, & Chong, 2004). More concretely, many of these models have assumed that individuals have or construct a set of strategies (Busemeyer & Myung, 1992; Erev & Barron, 2005). For example, Erev and Barron (2005) proposed a model to capture learning in DFE with partial-feedback and full-feedback. Reinforcement Learning Among Cognitive Strategies (RELACS) is a model that was initially proposed to account for learning effects in Roth and Erev (1995), and then further developed in Erev, Bereby-Meyer, and Roth (1999). RELACS assumes that the decision maker follows one of the three cognitive strategies or decision rules in each decision encountered and that the probability of using a strategy is determined by its frequency of use and the observed payoffs. RELACS has been shown to accurately replicate human choice behavior in different choice problems (Erev & Barron, 2005).

More recently, computational models based on IBLT (Gonzalez et al., 2003) have been shown to provide robust predictions in a variety of decision-making tasks in which individuals rely on experience (Lejarraga, Dutt, & Gonzalez, in press). IBLT was originally developed to explain and predict learning and decision making in real-time, dynamic decision-making environments. The theory has also been used as the basis for developing computational models that capture human decision-making behavior in a wider variety of DFE tasks. These include dynamically complex tasks (Gonzalez & Lebiere, 2005; Gonzalez et al., 2003; Martin, Gonzalez, & Lebiere, 2004), training paradigms of simple and complex tasks (Gonzalez, Best, Healy, Bourne, & Kole, 2010; Gonzalez & Dutt, 2010), simple stimulus–response practice and skill acquisition tasks (Dutt, Yamaguchi, Gonzalez, & Proctor, 2009), and repeated binary-choice tasks (Lebiere, Gonzalez, & Martin, 2007; Lejarraga et al., in press). The different applications of the theory illustrate its generality and ability to capture learning from exploration and DFE in multiple contexts.

Instance-Based Learning Theory

Instance-based learning theory uses a cognitive information representation called an instance. An instance consists of three parts: a situation (a set of attributes that define the decision situation), a decision or action, and an outcome or utility of the decision in the situation. The different parts of an instance are built through a general decision process with three steps: (a) characterizing a situation by attributes in the task, (b) making a decision based on expected utility, and (c) updating the utility in the feedback stage. The instances accumulated over time in memory are retrieved and used repeatedly in future decisions. Their strength in memory, called *activation*, is reinforced according to statistical procedures. These statistical procedures were originally developed by Lebiere (1999) and Anderson and Lebiere (1998) as part of the ACT-R cognitive architecture, which was used to build most of the existent IBL models (except for the more recent ones, such as in Lejarraga et al., in press).

In making a choice, IBLT recommends the selection of the alternative with the highest *blended value*, V (Lebiere, 1999), which is computed from all past instances where the outcome of that alternative was observed. The blended value of alternative j is defined as:

$$V_j = \sum_{i=1}^n p_i x_i \quad (1)$$

where x_i is the value of the observed outcome of alternative j in instance i , and p_i is the probability of retrieving instance i from memory. The blended value of an alternative is thus the sum of all observed outcomes x_i of corresponding instances in memory, weighted by their probability of their retrieval from memory. In any trial, the probability of retrieving instance i from memory is a function of that instance's activation relative to the activation of all other instances corresponding to that alternative, given by

$$p_i = \frac{e^{A_i/\tau}}{\sum_j e^{A_j/\tau}} \quad (2)$$

where τ is random noise defined as $= \sigma \times \sqrt{2}$, and σ is a free noise parameter. Noise in Equation 2 captures the imprecision of recalling instances from memory.

The activation of each instance in memory depends upon the *Activation* mechanism from ACT-R (Anderson & Lebiere, 1998):

$$A_i = \ln \left(\sum_{j=1}^n t_{ji}^{-d} \right) + \sum_{j=1}^m W_j \cdot (S - \ln(\text{fan}_{ji})) - \sum_{l=1}^k P_l \cdot M_{li} + N_i(0, \sigma) \quad (3)$$

The activation A_i of an instance i reflects the availability of an instance in memory. The activation is determined by the base-level activation $\ln \left(\sum_{j=1}^n t_{ji}^{-d} \right)$, the spreading activation $\sum_{j=1}^m W_j \cdot (S - \ln(\text{fan}_{ji}))$, the partial matching $\sum_{l=1}^k P_l \cdot M_{li}$, and noise $N_i(0, \sigma)$.

The base-level activation of an instance i reflects its recency and frequency of use in the past trials/times. n is the total number of previous uses of that instance from memory (frequency), including when the instance was first created, t_{ji} is the time elapsed since the j th use of the instance i (recency), and d is the memory's decay rate (a free parameter with a default value of 0.5). A high value of d (e.g., $d > 2.0$) means rapid decay of memory (forgetting easily) and reliance on recent past experiences, whereas a low value of d means slow decay of memory and reliance on temporally distant experiences.

The spreading activation reflects the impact of a task situation's attribute j on the instance i in memory. m is the total number of attributes in the situation, and W_j represents the amount of activation from attribute j to instance i . ACT-R assumes there is a limited total amount of attention W (a free parameter, with value of 1.0 as default), which gets divided equally among all the m attributes of the situation (thus, the W parameter is the sum of all W_j ; Anderson et al., 2004; Anderson, Reder, & Lebiere, 1996). Thus, the W parameter influences the attention to relevant and irrelevant attributes in the task's situation (Anderson et al., 1996; Lovett, Reder, & Lebiere, 1999). The term $(S - \ln(\text{fan}_{ji}))$ is called the strength of association. S (a free parameter that can be set to any value, whose default value of "nil" in ACT-R means not using spreading activation) is the maximum associative strength, and fan_{ji} is the number of instances in memory in which the attribute j appears as part of the instance's attributes plus one (as attribute j is associated with itself in instance i).

The partial matching component represents the mismatch between the attributes of a situation in the task and the attributes of the instance i in memory. k is the total number of attributes for a task's situation that are used to retrieve instance i from memory (thus, the summation is only computed over those attributes of an instance that correspond to the task's situation attributes used to retrieve the instance from memory). The match scale (P_l) reflects the amount of weighting given to the similarity between attribute l of an instance i in memory and the corresponding situation attribute in a task. P_l is generally defined to be a negative integer with a common value of -1.0 for all

attributes k of an instance i (but it could also be a free parameter with different negative real values for different attributes of an instance i). The M_{li} term, or match similarity, represents the similarity between the value l of an attribute in a task's situation (i.e., the situation attribute l from the task that is used to retrieve the instance i from memory) and the value in the corresponding attributes of the instance i in memory. Typically, M_{li} is defined using a linear similarity function that assumes M_{li} to be equal to the absolute value of the difference between a situation attribute from the task and the corresponding attribute value of an instance. Thus, the $\sum_{l=1}^k P_l \cdot M_{li}$ specification defines the similarity between the task's current situation attributes to the attributes of potentially retrievable instances from memory. Finally, the noise component is a Gaussian function with a mean of 0 and a standard deviation of σ .

Scaling up IBLT to Represent Effects of Social Interaction

In conflict situations, the social information that individuals have about their interaction partners can vary dramatically and thus influence both behavior and joint outcomes. Some conflict situations occur with minimal social information, as in cases where individuals affect one another's outcomes without even realizing their interdependence. The depletion of public goods is an example of low social information: if one person litters and contributes to pollution, another person living across town might suffer some incremental negative externality in terms of environmental quality, but the first person might not realize this impact, just as the second person might not be able to tie their poor outcome to the actions of the culprit. At the opposite end of the spectrum, some conflict situations occur with substantial levels of social information when individuals know one another's identities and motivations. An example of this is a husband who leaves dirty dishes on the counter for his wife to wash. Assuming that the two of them live alone together, they can attribute positive or negative actions to one another and put them in the context of their longer-term relationship.

Many models of individual DFE are incapable of representing response to social contexts. For example, Erev and Roth (2001) noted that simple reinforcement learning models predicted the effect of experience in two-person games like the IPD only in situations where players could not punish or reciprocate. A simple model predicts a decrease in cooperation over time even though most behavioral experiments demonstrate an increase in mutual cooperation because of the possibility of reciprocation (Rapoport & Chammah, 1965; Rapoport & Mowshowitz, 1966). To account for the effects of reciprocation, Erev and Roth used two explicit modifications to the basic model: if a player adopts a reciprocation strategy, he will cooperate in the next trial only if the other player has cooperated in the current trial; the probability that a player continues with the same strategy will depend on the number of times the reciprocation strategy was played (Erev & Roth, 2001). Although these tweaks to the model may accurately represent the kind of cognitive reasoning that people actually use in the IPD, they are unlikely to generalize to other situations with different action sets or outcomes.

Unlike models that require explicit modification to predict human behavior in conflict situations, we hypothesize that IBL models may offer the potential for generalization even

in their standard form, without predefined implementations of interaction strategies. Although most of the tasks modeled according to IBLT are individual tasks, there have been some initiatives to use IBLT in multiperson games. For example, Gonzalez and Lebiere (2005) reported a cognitive model of an IPD, initially reported by Lebiere, Wallach, and West (2000), which assumes instances are stored in memory, including one's own action, the other player's action, and the payoff. The decision is made according to a set of rules that, given each possible action, retrieves the most likely outcome from memory and selects the action with the highest payoff. This model was shown to accurately capture human behavior in the IPD. More recently, IBLT was also used as the basis for developing computational models that predict human behavior in more complex multiperson tasks, including a market entry game (Gonzalez, Dutt, & Lejarraga, in press). This model shares some basic features with IBL models of individual choice (e.g., Lejarraga et al., in press): weighting of prior outcomes by their probability of retrieval from memory (i.e., the blending mechanism); dependence on recency and frequency of past outcomes; and an inertia mechanism that depends upon surprise as a function of blended outcomes. In presenting a HSI that would influence social DFE, we propose some ways in which IBLT can account for more nuanced social aspects of interaction such as reciprocation, fairness, reputation, trust, and power.

Extensions of IBLT on a Hierarchy of Social Information

A basic HSI framework is shown in Figure 1, with each of the rows in the pyramid representing a category of interpersonal interaction characterized by the extent of social information possessed by involved parties. Regardless of the amount of social information, though, we assume that each individual knows their own actions, as well as seeing the outcomes of these actions over time. Each row in the pyramid builds upon the information available in the previous row. The differences between the categories reflect several key pieces of social information that can be studied separately to determine their incremental effects in naturalistic conflict situations. A more refined formulation could insert categories between those currently shown in Figure 1, as well as splitting the categories we have provided into further subcategories. We devised the taxonomy in its current form to offer sufficiently many levels for meaningful differentiation, but few enough for manageable comparison.

Nonsocial: Knowledge of One's Own Action-Outcome Pairs

The nonsocial category of information can be thought of as the individual behavior perspective, in which motivations for cooperative or competitive actions can be captured by reinforcement learning of two independent individuals in isolation. Nonetheless, even the actions that would maximize individual expected utility over time are not entirely straightforward because of the actual interdependence between individuals, which leads to shifting probabilities of various outcomes that may be perceived by decision makers as changes in the environment. Sometimes, these changes may seem completely unpredictable, but other times, decision makers may notice that their own actions influence the way that probabilities change. Thus, although individuals do not know that



Figure 1. Hierarchy of Social Information framework. Each level of the pyramid represents a category of interaction, with social information compounding from the bottom row (least information) to the top row (most information). Each interaction type includes all of the information below, as well as one new type of information.

another's behavior is affected by their own and vice versa, they may still observe that their present actions have an impact on their own future outcomes. In the pollution example mentioned earlier, one person tossing trash on the ground may lead another to conclude that there is no point in cleaning up after himself since the park remains filthy in either case and hence lead to a downward spiral of poor behavior. Alternately, if cleaning up contributes to a social norm of a clean park, then that behavior might be reinforced over time. In this sense, even if individuals were pure profit maximizers (against most evidence in the discussion to follow), there would still exist a tension between short-term gain (e.g., laziness for picking up trash) and long-term gain (e.g., a clean park where others pick up their trash too).

There have been some efforts to model optimal and actual human response to endogenous environmental changes in the absence of social concerns. For instance, Axelrod (1984) illustrated the purely rational tradeoff between short- and long-term profit maximization in a competition of computerized strategies playing the IPD against one another. The best performer on the objective criteria of total payoff was a simple strategy called tit-for-tat, which cooperates initially and then repeats its opponent's most recent action. Without "understanding" how it affects its opponent in any cognitive sense, tit-for-tat capitalizes on the possibility of mutual cooperation if the other is willing, but does not allow itself to be taken advantage of through asymmetrical cooperation. Axelrod characterizes the major benefits of tit-for-tat: it lacks a temptation to defect first, it retaliates against an opponent's defection, it is willing to forgive, and it is not eager to "beat" its opponent. Thus, while tit-for-tat indeed possesses some human-like characteristics, it may be more effective than humans who rarely follow axiomatic

principles so systematically. Baker and Rachlin (2001) showed that participants in an IPD adapt their behavior to reciprocation by a computerized opponent (which can be thought of a response from one's environment), but not in an optimal way. Specifically, people cooperated less than the optimal level when the computer's reciprocation probability was at an intermediate level and not made explicit. The authors described this as a failure to effectively trade off immediate payoff for greater payoff in future rounds (Baker & Rachlin, 2001).

Behavior in the nonsocial category is already well described by IBL models, as well as by other existent reinforcement learning algorithms. Given that one IBL model represents an individual, two identical versions of an IBL model would need to be created to act as a pair and interacting repeatedly in the case where two individuals interact. In the IPD, two identical models are created and decisions are made by using the mechanism described earlier. Interaction of these two models may occur under different informational conditions, which are expected to have different effects on the model's behavior. When the only information available is one's own actions and outcomes, the two models are expected to act independently, each representing instances with the action taken and the outcome obtained. The model parameters (e.g., d , σ) representing the recency of information and noise may act independently for each of the two simulated model participants in the IPD. Thus, the main prediction is that in the absence of any social information, IBL make decisions that depend exclusively on cognitive parameters at the individual level, including the frequency and recency of information from memory.

As we will see in the following sections, people may disobey strict payoff maximization in other ways when social information increases, and make it more difficult for a simple IBL model to capture these deviations.

Minimal: Knowledge of Human Interaction

With the exception of very conventional economists, most social scientists find it unsurprising that one fundamental type of social information available in an interaction—namely that one interacts with another human being—impacts the way people behave. Nonetheless, it is worth surveying research indicating the precise ways in which behavior with this knowledge has been shown to deviate from behavior without it. With regard to our theoretical framework, this is the distinction that we expect to observe in moving from the nonsocial category to the minimal category of interaction. Considerations of fairness and altruism on the positive side, as well as injustice or retribution on the negative, arise only in interaction where it is known that one has a human partner. For instance, Blount (1995) showed that playing with a human instead of a robot leads to more seemingly irrational rejections of low ultimatum game offers. While not a direct comparison of our nonsocial and minimal interaction categories where participants do not know the other's outcomes, these findings are suggestive that many noneconomic factors will come into play as soon as people believe they have some interdependence with others.

Differences in the nonsocial and minimal categories may be rooted in different neurological responses to other humans, as opposed to robots or random processes of nature. Several groups of researchers have shown differential brain activation when participants

believed they interacted with other humans versus computerized algorithms (Gallagher, Jack, Roepstorff, & Frith, 2002; McCabe, Houser, Ryan, Smith, & Trouard, 2001; Rilling, Sanfey, Aronson, Nystrom, & Cohen, 2004). These findings strongly suggest that one's own actions and outcomes are not the only variables encoded in the experience of events. When especially favorable or unfavorable outcomes are realized partly as a result of one's own choices, it will make a difference whether the unexplained variance can be attributed to the intentions of another person. The fundamental attribution error—defined as “the tendency for attributers to underestimate the impact of situational factors and to overestimate the role of dispositional factors in controlling behavior” (Ross, 1977, p. 183)—has been well documented in many social contexts. It may be particularly relevant here when the payoffs to the other person cannot be observed and put into perspective, allowing an individual's imagination to run wild with illusions that the other is achieving huge payoffs or even going out of their way to cause harm.

For example, in an empirical study of behavior in a variant of the IPD, Baker and Rachlin (2002) found that participants learned to cooperate at high rates when they knew they played against a probabilistic tit-for-tat algorithm, but not when they were led to believe that the algorithm was another human being. Although the false belief should have made no material difference to profit maximizing choices, the difference in cooperation can be accounted for by social considerations. It seems reasonable to speculate that participants fell prey to the attribution error when they were led to believe there was another human player (Ross, 1977). Since these participants knew nothing of the other's payoffs, it was easy to sustain an illusion that the other set out to cause harm ruthlessly or greedily. Believing that the other player was a human led to less cooperation because the downside risk of being taken advantage of was over-weighted relative to the possible upside of sustained cooperation. In another context, adding information about interdependence with others may have the reverse effect of enhanced cooperation.

The Minimal category of the HSI is expected to elicit different strengths and types of emotion than the nonsocial category because of players' awareness that their own actions affect other people and that other people's actions affect them. Although instance representations of an IBL model in the minimal category information category would be identical to that of the nonsocial category, the statistical mechanisms of the model are expected to have different effects in the model's predictions. We predict that a higher activation (Equation 3) would be necessary to highlight negative outcomes more than positive outcomes, because the negative outcomes would indicate that the counterpart defects on the PD. The effect of negative outcomes can be highlighted in IBL models through several mechanisms. First, given that blending includes the outcome directly and the selection of alternatives is based on those that have the highest blended value (see Equation 1), we would expect the model to be sensitive to the sign and magnitude of the outcome. Second, it is possible that people remember losses better than gains even in minimal social settings (e.g., recall instances of the other's defection better than instances of their cooperation). If this is the case, then the activation of instances with losses would be higher than the activation of instances with gains. The activations of those instances can be manipulated in several ways (see Equation 3); for example, by

reducing the value of the decay (d) parameter for losses, which will result in slow forgetting compared to gains.

Thus, the main prediction is that with knowledge of human interaction but absence of the other's outcomes and influence on one's outcomes, IBL decisions would be influenced by the emotional impact that positive or negative outcomes would have on our perception of the other's cooperation or defection. These effects could be captured by the explicit manipulation of the memory parameters that influence the memory for positive or negative outcomes.

Experiential: Knowledge of the Other's Actions and Outcomes

Taking as a given that an individual realizes their interdependence with another human being, there is evidence suggesting that behavior will also be affected by knowledge about the other person's particular actions and outcomes. In moving from the minimal social information category to the experiential social information category of the HSI framework, we expect to observe this distinction. This type of social information might encourage reasoning about the intentions behind the other's actions. On the one hand, this could further reduce the fundamental attribution error (Ross, 1977), whereby individuals might assume that the other has ill intentions rather than reasonable motives. Once the other's actions and payoffs are seen, it can be more easily acknowledged that the other faces a limited set of options and tradeoffs similar to one's own. Particularly, in cases of symmetric action sets and payoffs between two actors, as are most commonly used in the PD, the similarity of the other's predicament can be better related to one's own upon seeing their actions and outcomes. This type of empathy could lead to greater cooperation in such interactions.

On the other hand, people may be especially sensitive to asymmetrically unfavorable outcomes when they are able to directly compare their own outcomes with those of another person. For example, Silverstein et al. (1998) found that making choices for four sequential trials before learning the other's actions impeded tendencies toward mutual defection in an IPD. The authors suggest that immediate discovery of the other's action can cause people to succumb to the fear or impulse to defect more so than when they only discover the other's actions after several of their own decisions have been made. With directionally similar results in another setting, Erev and Rapoport (1998) showed in a laboratory study that entry into a competitive market increased relative to the social optimum when players saw others' payoffs. This is consistent with an account where individuals' reference points for their own payoffs are influenced by the payoffs of others. A desire for equity can also explain in large part why people consistently offer more than the minimal amount to others in the ultimatum game, and why low offers are rejected (Fehr & Schmidt, 1999). These authors conclude that either fair or selfish types can prevail, depending on the strategic environment and opportunity for some people to punish others with antisocial behavior. Despite the contrasting empirical evidence that information about others' actions and payoffs can lead to either more or less cooperation, this type of information will generally induce a greater sense of accountability since it is more difficult to misconstrue one another's intentions.

The experiential category of the HSI framework will be more challenging for IBL to model. As before, two identical versions of an IBL model would be created to act as a pair and interacting repeatedly. Given that this layer includes social comparisons and fairness concerns; however, we expect that the representation of information in an instance will need to be changed. For example, individuals involved in conflict might explicitly represent the outcomes of others in addition to their own experienced outcomes. The question of how these outcomes are represented and how they are stored in memory is an open question. It is possible that the “Situation” in an instance would include not only each person’s choice and outcome, but also the other person’s choice and outcome. This would create a dependency as the content of own and other’s actions and outcomes are different for each model. It is also possible that the utility represented in an instance is formed by the *comparison* of the outcomes experienced; the difference between the own and the other’s outcome for example. If so, individuals may have a nonlinear representation for the difference between their own outcome and the outcome of another person, likely with a discontinuity at zero and steeper slope below zero. It is also cognitively plausible that the decay of instances in memory (determined by the d parameter) depends on these relative outcomes, so this presents another candidate for an IBL model extension that may be required to capture the experiential condition.

Descriptive: Knowledge of the Complete Interaction Structure

Another distinction that should be largely immaterial to behavior from at least an economic standpoint is the way that people discover their own and another’s outcomes. Given a relatively small set of possible actions, the long-run behavior should be similar whether or not these outcomes were described from the outset of interaction if people eventually experience all possible outcomes. Yet, the reason for including this difference in our HSI framework is a distinction between the experiential and descriptive categories supported by existing research showing that the discovery process does influence interactive behavior. Most notably, whether or not individuals possess a full description of the interaction structure can affect the subjective impact of outcomes to them, and therefore, preferences for some actions over others. For example, Rapoport, Guyer, and Gordon (1976) demonstrated greater levels of overall cooperation in an IPD where a complete payoff matrix was displayed to participants than when they simply learned which outcomes corresponded to each set of actions by themselves and their opponents through trial-and-error.

Furthermore, we argue that ongoing visibility of the interaction structure can assist in putting briefly painful or joyful outcomes in the broader context of mutual interdependence. Thus, compared to the experiential interaction category where recent outcomes tend to generate a strong emotional impact, reactions to both positive and negative outcomes could be moderated in the descriptive category where the outcome is seen as merely one encounter within a long-term relationship. Particularly in symmetric games where individuals can see that others face a dilemma similar to their own, the presence of descriptive social information can aid in taking the other’s perspective and behaving as one would wish the other to behave. Beyond this, descriptive reminders may induce an individual to engage in second-order reasoning about how their own actions will be

perceived by others. The sense that one's counterpart is reacting to one's own actions may, in itself, lead to a feeling of empowerment that voluntarily directs the course of interaction toward mutual cooperation rather than conflict escalation.

The descriptive category, where the sets of outcomes from all actions are known ahead of time and observed continually, will require different assumptions in the IBL model. For example, it is possible that initial instances including descriptive information of the outcomes are created in human memory. These initial instances could be inserted into the model to create human-like expectations of long-term dependencies. An IBL model may also require the inclusion of foregone payoffs in the utility calculation to capture the availability of this information in the descriptive category. It is difficult to measure in laboratory experiments and therefore will be difficult to model whether people may also compare their experienced outcomes with foregone outcomes or instead treat the foregone outcomes as if they had been experienced from making the opposite choice. More generally, we expect that people engaged in descriptive interaction will be able to reason somewhat more strategically about the extended consequences of their own and others' behavior. If this is so, it will be reflected in behavior (e.g., higher levels of cooperation in the IPD), which will only be adequately fit by a variation of the IBL model that similarly incorporates descriptive social information.

Cultural: Knowledge of the Other's Identity

The identity of one's interaction partner is not a neutral piece of information, either from a strategic or an emotional standpoint. In moving from the descriptive category to the cultural category of interaction in our HSI framework, several dimensions of identity might be strategically relevant. The mere knowledge of demographic characteristics or the group membership of another person might shape expectations about their likelihood of engaging in certain actions. Early work on cross-cultural interaction, for instance, tended to focus on classifying differences in attitudes, beliefs, and values across cultures (Hofstede, 1983). With much effort dedicated to this line of inquiry, there have been some differences in interaction styles established on the basis of observable characteristics like gender (Babcock & Laschever, 2003) and country of origin (Buchan, Croson, & Dawes, 2002). Of course, unless individuals are personally acquainted, there is a strong likelihood of drawing false conclusions about inherent behavioral proclivities based on stereotypes. Perhaps, more common than actual persistent differences in interaction tendencies between groups is the triggering or situational activation of group stereotypes that lead to a self-fulfilling prophesy in behavior (Kray, Thompson, & Galinsky, 2001; Riley Bowles, Babcock, & McGinn, 2005). In fact, more fundamental individual differences in social motive and punitive capability may be unrelated to group membership (De Dreu, Giebels, & Van de Vliet, 1998). Power dynamics may also come into play if individuals recognize their status with respect to others. Galinsky, Magee, Inesi, and Gruenfeld (2006) observed that greater power can impede perspective taking and hence may lead individuals to ignore the harm or benefit they cause to others.

More important than identity information that has some *de facto* strategic relevance to predicting others' behavior, the identity of another person might affect the degree to which one feels altruism or aggression toward him. An extensive literature details

in-group favoritism or the tendency to bestow disproportionate rewards on similar others and disproportionate punishments on dissimilar others. There may be an evolutionary basis for in-group bias: at one point in human history, one's own survival was closely intertwined with that of kin group members (Trivers, 1971). If in-group favoritism has been biologically hard wired, it might manifest as favoritism and empathy toward those who share interests other than survival. Decreasing social distance between interaction partners, for example, leads to more generous offers in the dictator game, arising from social norms of reciprocity (Hoffman, McCabe, & Smith, 1996) or from "other-regardness independent of any norms of social exchange" (Bohnet & Frey, 2006, p. 335). Such parochial altruism is a robust finding from laboratory studies (e.g., Intergroup PD experiments in Halevy, Bornstein, & Sagiv, 2008), naturally occurring groups (e.g., indigenous tribes in Papua New Guinea documented by Bernhard, Fischbacher, & Fehr, 2006), and randomized groups that develop strong identifications (e.g., Swiss army platoons in Goette, Huffman, & Meier, 2006). In sum, knowing the identity of an interaction partner may affect one's own actions based on logical inferences about what actions the other will take, as well as altering one's own actions based on altruistic or punitive motives directed at the other.

Based on the important behavioral distinctions between the cultural category and those below it in the HSI framework, this may prove to be the most challenging layer of social information to translate into IBLT. For instance, attention and memory may be triggered differently for individuals depending on the identity traits of their interaction partners relative to their own. This can be incorporated into an IBL model through the similarity metrics in Equation 3. Previous instances in which people interacted with others similar to a current interaction partner will be more strongly activated. Sometimes, this might lead to the formation of stereotypes where it is expected that people will behave like others with whom they share observable identity traits. Another way to include the effects of in-group favoritism in an IBL model is to capture the way individuals experience the outcomes of others similar or dissimilar to themselves. The outcomes of another can be included as outcomes for oneself in an instance, as if the individual feels a counterpart's joy or pain. At this stage, these are merely speculations about how IBLT—which has proven to be versatile for a range of individual and several multiplayer decision tasks—may be more extensively applied to modeling the impacts of various levels of social information that we have laid out in the HSI framework.

Conclusions

Given the complexities of real-world conflict that cannot be completely captured by other cognitive models, we propose IBLT as one promising new approach. DFE have rarely been studied in social contexts (Erev & Haruvy, in press), and this new approach and shift from the classical decisions from description will bring new light onto our understanding of general conflict resolution research. Because models generated using IBLT have been successful in describing DFE in a variety of tasks, we argue that IBLT will also offer flexibility for representing nuances of social information. In individual choice tasks, IBL models are able to fit human behavior without depending on assumptions about either

rationality (i.e., expected utility maximization) or interpretations of descriptive information (e.g., weighting of probabilities in a fully defined gamble). Instead, the model's decisions emerge naturally and implicitly from the way instances are encoded and retrieved from memory, similar to the way that we believe people process information in the real world. Our objective in using IBLT here is to push its capabilities a step further to encompass social information that decision makers encounter in interaction with others.

The Hierarchy of Social Information aims to characterize the levels of information that have been shown to impact behavior in conflict situations and thus puts forth a clear challenge for the behaviors that we would like IBL models to encompass. The HSI framework assumes that individual actions and outcomes within social interaction are influenced by the amount and type of information available to decision makers. The distinct categories range from one in which individuals possess no information about their interdependence with others, to one in which individuals have complete information about the actions of others, their influence on each other's outcomes, and the identities of others. A large body of existing research indicates that behavior differs significantly in moving from one category to the next, and yet these differences have not been systematically tested in either experiments or cognitive models. We discuss how these informational characteristics may influence reciprocity, fairness, trust, power, and other social considerations, especially as individuals are better able to perceive the intentions of others and to reason about how their own intentions might be perceived. By offering a clear separation of varying degrees of social information, the HSI framework allows the formulation of hypotheses about how cognitive and emotional processing will shift as higher degrees of information become available, how this information will induce behavioral differences in the laboratory, and how this information can be represented in IBL models without altering their defining components.

Describing and predicting social behavior computationally is a current research challenge with many unanswered questions: How can we account for nonmonetary incentives such as fairness, trust, reciprocity, and power? Do social effects need to be explicitly represented in models that correctly predict behavior? How do individual cognitive constraints influence the use of social information? How do we scale our behavioral measures and models from the individual level to the social level? The HSI framework we present can help to organize research comparing decision making across informational categories, so that we have a clear set of behavioral results that we would like to better understand. In addition, we propose that IBLT might offer a fruitful approach to modeling empirical differences in conflicts across the HSI, based on the robustness of IBL models for fitting individual choice tasks and the ease of generalizing core mechanisms to new situations.

References

- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*(4), 1036–1060.
- Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Hillsdale, NJ: Lawrence Erlbaum Associates.

- Anderson, J. R., Reder, L. M., & Lebiere, C. (1996). Working memory: Activation limitations on retrieval. *Cognitive Psychology*, *30*(3), 221–256.
- Axelrod, R. (1980). Effective choice in the Prisoner's Dilemma. *Journal of Conflict Resolution*, *24*(1), 3–25.
- Axelrod, R. (1984). *The evolution of cooperation*. New York, NY: Basic Books.
- Babcock, L., & Laschever, S. (2003). *Women don't ask: Negotiation and the gender divide*. Princeton, NJ: Princeton University Press.
- Baker, F., & Rachlin, H. (2001). Probability of reciprocation in repeated prisoner's dilemma games. *Journal of Behavioral Decision Making*, *14*, 51–67.
- Baker, F., & Rachlin, H. (2002). Teaching and learning in a probabilistic prisoner's dilemma. *Behavioural Processes*, *57*, 211–226.
- Bernhard, H., Fischbacher, U., & Fehr, E. (2006). Parochial altruism in humans. *Nature*, *442*(7105), 912–915.
- Blount, S. (1995). When social outcomes aren't fair. The effect of causal attributions on preferences. *Organizational Behavior and Human Decision Processes*, *63*, 131–144.
- Bohnet, I., & Frey, B. S. (2006). Social distance and other-regarding behavior in dictator games: Comment. *American Economic Review*, *89*(1), 335–339.
- Buchan, N. R., Croson, T. A., & Dawes, R. M. (2002). Swift neighbors and persistent strangers: A cross-cultural investigation of trust and reciprocity in social exchange. *American Journal of Sociology*, *108*(1), 168–206.
- Busemeyer, J. R., & Myung, I. J. (1992). An adaptive approach to human decision making: Learning theory, decision theory, and human performance. *Journal of Experimental Psychology: General*, *121*(2), 177–194.
- Bush, R. R., & Mosteller, F. (1955). *Stochastic models for learning*. Oxford, U.K.: John Wiley & Sons, Inc.
- Colman, A. M. (2003). Cooperation, psychological game theory, and limitations of rationality in social interaction. *Behavioral and Brain Sciences*, *26*, 139–198.
- De Dreu, C. K. W., Giebels, E., & Van de Vliet, E. (1998). Social motives and trust in integrative negotiation: The disruptive effects of punitive capability. *Journal of Applied Psychology*, *83*(3), 408–422.
- Dutt, V., Yamaguchi, M., Gonzalez, C., & Proctor, R. W. (2009). An instance-based learning model of stimulus-response compatibility effects in mixed location-relevant and location-irrelevant tasks. In A. Howes, D. Peebles & R. Cooper (Eds.), *9th International Conference on Cognitive Modeling—ICCM2*. Manchester, U.K.
- Erev, I., & Barron, G. (2005). On adaptation, maximization, and reinforcement learning among cognitive strategies. *Psychological Review*, *112*(4), 912–931.
- Erev, I., Bereby-Meyer, Y., & Roth, A. E. (1999). The effect of adding a constant to all payoffs: Experimental investigation, and implications for reinforcement learning models. *Journal of Economic Behavior & Organization*, *39*, 111–128.
- Erev, I., Glözman, I., & Hertwig, R. (2008). What impacts the impact of rare events. *Journal of Risk and Uncertainty*, *36*(2), 153–177.
- Erev, I., & Haruvy, E. (in press). Learning and the economics of small decisions. In J. H. Kagel & A. E. Roth (Eds.), *The handbook of experimental economics* (Vol. 2). Princeton, NJ: University Press.
- Erev, I., & Rapoport, A. (1998). Coordination, “magic,” and reinforcement learning in a market entry game. *Games and Economic Behavior*, *23*, 146–175.

- Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *The American Economic Review*, 88(4), 848–881.
- Erev, I., & Roth, A. E. (2001). Simple reinforcement learning models and reciprocity in the Prisoner's Dilemma game. In G. Gigerenzer & R. Selten (Eds.), *Bounded rationality: The adaptive toolbox* (pp. 215–231). Cambridge, MA: MIT Press.
- Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition and cooperation. *Quarterly Journal of Economics*, 114(3), 817–868.
- Galinsky, A. D., Magee, J. C., Inesi, M. E., & Gruenfeld, D. H. (2006). Power and perspectives not taken. *Psychological Science*, 17(12), 1068–1074.
- Gallagher, H., Jack, A., Roepstorff, A., & Frith, C. D. (2002). Imaging the intentional stance in a competitive game. *NeuroImage*, 16, 814–821.
- Goette, L., Huffman, D., & Meier, S. (2006). The impact of group membership on cooperation and norm enforcement: Evidence using random assignment to real social groups. *American Economic Review*, 96(2), 212–216.
- Gonzalez, C., Best, B. J., Healy, A. F., Bourne, L. E. Jr, & Kole, J. A. (2010). A cognitive modeling account of simultaneous learning and fatigue effects. *Journal of Cognitive Systems Research*, 12, 19–32.
- Gonzalez, C., & Dutt, V. (2010). Instance-based learning models of training. In *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting* (pp. 2319–2323). San Francisco, CA: Human Factors and Ergonomics Society.
- Gonzalez, C., Dutt, V., & Lejarraga, T. (in press). A loser can be a winner: Comparison of two instance-based learning models in a market entry competition. *Games*.
- Gonzalez, C., & Lebiere, C. (2005). Instance-based cognitive models of decision making. In D. Zizzo & A. Courakis (Eds.), *Transfer of knowledge in economic decision-making* (pp. 148–165). New York: Macmillan (Palgrave Macmillan).
- Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science*, 27(4), 591–635.
- Halevy, N., Bornstein, G., & Sagiv, L. (2008). 'In-group love' and 'out-group hate' as motives for individual participation in intergroup conflict—A new game paradigm. *Psychological Science*, 19(4), 405–411.
- Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, 15(8), 534–539.
- Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2006). The role of information sampling in risky choice. In K. Fiedler & P. Juslin (Eds.), *Information sampling and adaptive cognition* (pp. 72–91). New York: Cambridge University Press.
- Ho, T. H., Camerer, C. F., & Chong, J. K. (2004). *The economics of learning models: A self-tuning theory of learning in games*. Berkeley: University of California.
- Hoffman, E., McCabe, K., & Smith, V. L. (1996). Social distance and other-regarding behavior in dictator games. *American Economic Review*, 86(3), 653–660.
- Hofstede, G. (1983). National cultures in four dimensions: A research-based theory of cultural differences among nations. *International Studies of Management and Organization*, 13, 46–74.
- Kray, L. J., Thompson, L., & Galinsky, A. (2001). Battle of the sexes: Gender stereotype activation in negotiations. *Journal of Personality and Social Psychology*, 80, 942–958.

- Lebiere, C. (1999). Blending: An ACT-R mechanism for aggregate retrievals. Paper presented at the Sixth Annual ACT-R Workshop at George Mason University.
- Lebiere, C., Gonzalez, C., & Martin, M. (2007). Instance-based decision making model of repeated binary choice. In R. L. Lewis, T. A. Polk & J. E. Laird (Eds.), *Proceedings of the 8th International Conference on Cognitive Modeling* (pp. 67–72). Ann Arbor, MI: Taylor & Francis/ Psychology Press.
- Lebiere, C., Wallach, D., & West, R. L. (2000). A memory-based account of the prisoner's dilemma and other 2x2 games. Paper presented at the International Conference on Cognitive Modeling.
- Lejarraga, T., Dutt, V., & Gonzalez, C. (in press). Instance-based learning: A general model of repeated binary choice. *Journal of Behavioral Decision Making*.
- Lovett, M. C., Reder, L. M., & Lebiere, C. (1999). Modeling working memory in a unified architecture: An ACT-R perspective. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 135–182). New York: Cambridge University Press.
- March, J. G. (1996). Learning to be risk averse. *Psychological Review*, 103(2), 309–319.
- Martin, M. K., Gonzalez, C., & Lebiere, C. (2004). Learning to make decisions in dynamic environments: ACT-R plays the beer game. In M. C. Lovett, C. D. Schunn, C. Lebiere & P. Munro (Eds.), *Proceedings of the Sixth International Conference on Cognitive Modeling* (Vol. 420, pp. 178–183). Pittsburgh, PA: Lawrence Erlbaum Associates Publishers.
- McCabe, K., Houser, D., Ryan, L., Smith, V., & Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Academy of Sciences*, 98, 11832–11835.
- Rabin, M. (1993). Incorporating fairness into game theory and economics. *The American Economic Review*, 83(5), 1281–1302.
- Rapoport, A., & Chamah, A. M. (1965). *Prisoner's dilemma: A study in conflict and cooperation*. Ann Arbor: University of Michigan Press.
- Rapoport, A., Guyer, M. J., & Gordon, D. G. (1976). *The 2x2 game*. Ann Arbor: University of Michigan Press.
- Rapoport, A., & Mowshowitz, A. (1966). Experimental studies of stochastic models for the Prisoner's dilemma. *System Research and Behavioral Science*, 11(6), 444–458.
- Riley Bowles, H., Babcock, L., & McGinn, K. (2005). Constraints and triggers: Situational mechanics of gender in negotiation. *KSG Working Paper*, No. RWP05-051.
- Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2004). The neural correlates of theory of mind within interpersonal interactions. *Neuroimage*, 22, 1694–1703.
- Roberts, G. (1997). Testing mutualism: A commentary on Clements and Stephens. *Animal Behaviour*, 53(6), 1361–1362.
- Ross, L. D. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed), *Advances in experimental social psychology* (Vol. 10, pp. 173–220). New York: Academic Press.
- Roth, A. E., & Erev, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior*, 8, 164–212.
- Schuster, R. (2000). How useful is an individual perspective for explaining the control of social behavior? *Behavioral and Brain Sciences*, 23(2), 263–264.

- Schuster, R., & Perelberg, A. (2004). Why cooperate? An economic perspective is not enough. *Behavioural Processes*, 66, 261–277.
- Silverstein, A., Cross, D., Brown, J., & Rachlin, H. (1998). Prior experience and patterning in a prisoner's dilemma game. *Journal of Behavioral Decision Making*, 11, 123–138.
- Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, 46(1), 35–57.
-

Cleotilde Gonzalez is an Associate Research Professor and Director of the Dynamic Decision Making Laboratory (<http://www.cmu.edu/ddmlab>), Department of Social and Decision Sciences, Carnegie Mellon University. She is an affiliated faculty at the Human–Computer Interaction Institute, Center for Cognitive Brain Imaging, and Center for Neural Basis of Cognition at Carnegie Mellon, and the Center for Research on Training at the University of Colorado. She is in the Editorial board of the Human Factors Journal and Associate Editor of the Journal of Cognitive Engineering and Decision Making.

Jolie Martin is a Postdoctoral Researcher in the Dynamic Decision Making Laboratory at Carnegie Mellon University. She received her PhD in Science, Technology, and Management (2008) and MBA (2005) from Harvard Business School. She also served as an Adjunct Lecturer in Negotiation at the Harvard Kennedy School of Government and Postdoctoral Researcher in Trust, Emotion, Ethics, and Morality at Harvard Law School (2008–09). Her current research interests include conflict resolution, consumer behavior, cultural identification, decision making under uncertainty, experimental economics, game theory, and negotiation.