Taming Uncertainty

Ralph Hertwig, Timothy J. Pleskac, Thorsten Pachur,
and The Center for Adaptive Rationality

The MIT Press
Cambridge, Massachusetts
London, England

5 Strategic Uncertainty and Incomplete Information: The Homo Heuristicus Does Not Fold

Leonidas Spiliopoulos and Ralph Hertwig

5.1 Strategic and Environmental Uncertainty

In the movie *The Hunt for Red October*, based on the novel by Tom Clancy (1984) and set at the height of the Cold War, Soviet captain Marko Ramius has been given command of a prototype nuclear submarine with a revolutionary stealth propulsion system (Neufeld & McTiernan, 1990). *Red October's* new technology would allow it to approach the U.S. coast undetected and initiate a first strike with little to no warning. Understanding the consequences of this imbalance between the two superpowers, Captain Ramius plans to defect to the United States, gifting to it the stealth technology and thus eliminating the Soviets' strategic advantage. As Ramius is heading toward U.S. waters, the Soviets learn that he intends to defect and deploy other submarines to stop him. The United States is now in a precarious position. Soviet diplomats tell their U.S. counterparts that Captain Ramius has gone rogue and intends to launch an unauthorized nuclear strike on the United States; they ask for help in capturing him or destroying *Red October*. The U.S. command brings in a CIA analyst, Jack Ryan, to assess the situation. Ryan tries to infer Ramius's motivation, state of mind, and intentions: Is he a rogue captain capable of launching a nuclear attack? Or does he have some other (as yet unknown) motivation? Is he insane or sane? Drawing on prior information about Captain Ramius and interpreting his actions during the rapidly unfolding events, Ryan reaches the correct conclusion: Ramius intends to defect rather than to strike.

Uncertainty about an opponent's characteristics—in terms of preferences, beliefs, or strategies—is termed *strategic uncertainty*. Ramius knows his own intentions and preferences, but they are not common knowledge. If the United States were persuaded by the Soviets' framing of the situation, their best response—given their limited knowledge of the situation—would be to help sink the vessel. In the long run, this would be to their disadvantage, as they would lose access to the game-changing stealth technology. But as the U.S. command is unaware of *Red October*'s new technology, it cannot properly evaluate the consequences of its actions. Thus, not only are the U.S. decision makers exposed to strategic uncertainty about their opponent, but they are also uncertain or ignorant of the payoffs of their actions (*environmental uncertainty*; see also chapter 12). The interaction of strategic uncertainty and incomplete information makes this game a particularly difficult and unpredictable one—not least because it has no historical precedent. In game theory, this "ahistorical" interaction is called a one-shot game—in contrast to a repeated game, where the same game may be played against the same opponent multiple times.

How can people resolve, reduce, or otherwise handle the various kinds of uncertainty involved in strategic one-shot interactions—that is, interactions that do not afford the opportunity for learning? Is this domain one in which *Homo heuristicus*, equipped with computationally modest and informationally frugal decision rules, will thrive or falter? Using computer simulations, we compared the performance of random behavior (a baseline), eight heuristics, and a normative solution (the Nash equilibrium) across environments with different levels of strategic and environmental uncertainty. We refer to these collectively as decision policies. Our focus is on heuristics for two key reasons: First, computation of the normative solution becomes so complex as to be psychologically implausible once strategic and/or environmental uncertainty are entered in the equation. Second, simple heuristics can be as accurate as, and sometimes even more accurate than, decision policies that take all the available information into account, including optimization models (Gigerenzer & Brighton, 2009; see chapter 2). This finding—that simplicity does not necessarily come at the cost of accuracy—has raised a crucial question: In which environments can a simple heuristic outperform optimizing policies, and in which will it lag behind? In discussing the "right" environments for heuristics, the philosopher Sterelny (2003) has suggested that they are likely to succeed only in nonsocial environments, that is, in environments that rarely involve "competitive, interacting, responsive aspects of the environment" (p. 208): "For it is precisely in such situations that simple rules of thumb will go wrong.... Catching a ball is one problem; catching a liar is another" (p. 53).

Our goal is to examine precisely this question: Do heuristics fail in the face of strategic uncertainty and incomplete information? We further explore to what extent heuristics' performance in strategic games is contingent on characteristics of the environment—possibly suggesting a map of bounded rationality that describes relationships between regions (properties) of the environment and properties of heuristics that can either boost or undermine the accuracy of decisions.

5.2 The Unwieldy Normative Solution in Strategic Games

Let us first consider the normative solution. A set of strategies constitutes a Nash equilibrium (see box 5.1 for an example) if no player can gain (increase their payoff) by unilaterally changing their strategy. It is assumed that each player's beliefs about their opponent's action are correct, that players are fully rational—that is, they best respond to their beliefs—and that each player is aware of the other player's strategies (common knowledge assumption). The crucial construct that clears up strategic uncertainty at the equilibrium is the assumption that both players' beliefs about their opponent's behavior are correct.

In games of incomplete information with strategic and environmental uncertainty, a refinement of the standard equilibrium solution, the Bayesian Nash equilibrium (Harsanyi, 1968), assumes correct (probabilistic) beliefs about the possible states of uncertain variables (e.g., payoffs, others' beliefs). In the *Red October* example, Captain Ramius can be modeled as either insane or sane, with the respective state determining his payoffs (assigning high positive or negative utility, respectively, to initiating a nuclear strike). The Bayesian Nash equilibrium requires that people hold consistent beliefs about the probability distribution of their opponent's preferences (e.g., Ramius is insane with $p=0.6$ and sane with $p=0.4$). Similarly, the Bayesian Nash equilibrium requires consistent beliefs about the distribution of the uncertain environmental payoffs (e.g., the payoff is equal to x with $p=0.3$ or to y with $p=0.7$). Uncertainty is resolved through the forced consistency of beliefs at the equilibrium, and the rationality criterion is met by the choice of actions that represent the best responses given those beliefs.

Typically, as illustrated in the example above, beliefs are modeled as exact probability estimates; in other words, the uncertainty associated with the incomplete information is assumed to be reducible to risk (a known probability). But what if it is not possible to assign exact probabilities to the likelihood that Ramius is sane or insane? Extending the Bayesian Nash equilibrium to uncertainty makes it even more complex. In fact, it renders the solution practically intractable, as it would require second-order probability distributions for all the uncertain variables. Note that the indiscriminate application of a Bayesian approach to strategic decision making under uncertainty is widely contested, as are the Bayesian Nash equilibrium's demanding assumptions, even by prominent game theorists (e.g., Binmore, 2007).

Another complication with the normative solution is that multiple Nash equilibria may exist in a single game. Strategic uncertainty cannot be completely resolved in games with multiple equilibria, even under the strict rationality postulates of the Nash equilibrium. Various theories of equilibrium selection address this uncertainty by whittling down the set of Nash equilibria to a single one: payoff- versus risk-dominance equilibrium (Harsanyi & Selten, 1988), trembling-hand equilibrium (Selten, 1975), or evolutionary selection (Kandori, Mailath, & Rob, 1993). However, all of these modifications require ever more complex rationalizations and beliefs about how the opponent selects the equilibrium. They are therefore unlikely to be good, psychologically plausible descriptive models of human behavior—at least in one-shot games.

In repeated games, where players can accumulate experience with the same game and counterpart, it is possible that they inductively learn and converge to the Nash equilibrium (Binmore, Swierzbinski, & Proulx, 2001; Erev & Roth, 1998; Ochs, 1995; Roth & Erev, 1995). But what happens in uncertain environments comprised of one-shot games against different players? In these conditions, it is difficult or impossible to reduce uncertainty by inductive learning. Are players able to deductively solve the Nash equilibrium? The empirical findings suggest that most participants do not play according to the Nash equilibrium. Instead, they appear to employ various boundedly rational decision rules that we refer to as heuristics.

5.3 Why Do People Resort to Simple Heuristics in Strategic Games?

There are at least four answers to this question. First, people are simply unlikely to be able to calculate the Nash equilibrium. The implication here is that they should use the Nash equilibrium (see box 5.1 for an illustration)—but that in reality, as boundedly rational decision makers, they lack the cognitive capacity to do so (Kahneman, 2003b). Second, along with high mental effort, computing the Nash equilibrium exacts high decision costs. A perfectly rational player should take these costs into account and optimize with respect to the time, computation, money, and other resources incurred (e.g., Sargent, 1993). If the decision costs are high enough, it may be rational to save resources—for instance, by resorting to a simpler decision rule. Third, the Nash equilibrium is a rational response only if a player believes that their opponent will also behave rationally. If a rational player comes to believe that their opponent will, for whatever reason, deviate from the Nash equilibrium, they may also switch gears.

Finally, we propose a fourth possibility that is not necessarily incompatible with some of the previous ones. In uncertain environments, heuristics often perform on par with or even better than more complex decision rules. This has been repeatedly demonstrated in games against nature (e.g., Gigerenzer, Todd, & the ABC Research Group, 1999; Hertwig, Hoffrage, & the ABC Research Group, 2013) but less so in games against others. Our goal is to show that what holds for nonsocial environments, in which a person plays against disinterested nature, also holds in environments involving other people. Social environments have often been characterized as more challenging and intellectually demanding than nonsocial ones (see Hertwig & Hoffrage, 2013). Yet, even here—or perhaps particularly here—simple heuristics may be as accurate as, and sometimes even more accurate than, policies that make the greatest possible use of information and computation. We examined how several simple heuristics fare relative to the normative solution in strategic one-shot games. We studied the heuristics' *ecological rationality*, seeking to identify the environmental properties that are conducive to good performance. Our approach parallels investigations into cue-based inference (see Hogarth & Karelaia, 2006). Our aim was to propose a first map of bounded rationality in strategic games, revealing properties of the games being played, and determining the degree of environmental and strategic uncertainty under which heuristics perform well—or fall behind.

Box 5.1
The Nash equilibrium and strategic dominance.

The Nash equilibrium of the game presented in table 5.1 is the combination of player A's Down strategy and player B's Left strategy. It is calculated as follows: for each combination of possible strategies, ask the question whether either player has an incentive to change their strategy (assuming that the opponent's strategy remains the same). If the answer for both players is no, then the combination of strategies constitutes a Nash equilibrium. Let us start in the upper left cell of table 5.1: Up, Left. If player B chooses Left, does player A have an incentive to switch from Up to another strategy? The answer is yes: the payoff for Up is 58 units, but the payoff for Down is 70 units. Does player B have an incentive to switch from Left (with a payoff of 57) to another strategy? Yes, playing Center would give player B a payoff of 89. Therefore, the combination Up, Left is not a Nash equilibrium. The same reasoning can be applied to every combination of strategies. Let us skip directly to the combination Down, Left. Player A achieves a payoff of 70, the maximum available if player B chooses Left. Player B receives a payoff of 74, the maximum available if player A chooses Down. As neither player has an incentive to change, this combination of strategies constitutes a Nash equilibrium.

Table 5.1 also illustrates strategic dominance. We begin with self-dominance. Up dominates Middle for player A, because the payoffs for Up are always larger than those of Middle, whichever strategy player B chooses. If player B chooses Left, player A receives a payoff of 58 for Up relative to 34 for Middle. If player B chooses Center, player A receives 32 for Up relative to 23 for Middle. If player B chooses Right, player A receives 94 for Up and 37 for Middle. Therefore, Middle is a dominated strategy. Analogous comparisons can be performed to determine opponent-dominance.

5.4 Testing Simple Heuristics in Strategic Games

Herbert Simon (1955, 1956) saw bounded rationality in terms of two interlocking components: the limitations of the human mind and the structure of the environment. The implications of this conceptualization are twofold. First, models of simple heuristics need to reflect the mind's actual capacities rather than assume it to have unbounded capabilities. Second, the environmental structure may be the key to a heuristic's performance, to the extent that the heuristic's architecture successfully maps onto the structure of the environment (or parts of it). For this reason, we now describe our computer simulations in terms of the game environment and the competing heuristics.

Table 5.1
An example of a 3×3 normal-form game.

	Player B		
Player A	Left	Center	Right
Up	58, 57	32, 89	94, 41
Middle	34, 46	23, 31	37, 16
Down	70, 74	41, 12	23, 53

5.4.1 The Strategic Games

The games are all one-shot normal-form games with simultaneous moves, meaning that players make their decisions simultaneously—and therefore independently, without first being able to observe the opponent's choice. Table 5.1 gives an example of a normal-form game with two players. Player A can choose one of the strategies in the rows (Up, Middle, Down); player B can choose one of the strategies in the columns (Left, Center, Right). The number of possible actions for each player is denoted by n—we assume that it is the same for both players and refer to this as the *size* of the game. The game in table 5.1 is therefore a 3×3 game, with each player having three strategies at their disposal. The numbers represent the payoffs that each player will receive for every possible combination of strategies. The first number represents player A's payoff; the second, player B's payoff. For example, if player A chooses Up and player B chooses Center, they receive 32 and 89 units, respectively.

The set of strategic games we investigated included, but was not restricted to, widely researched 2×2 games such as the prisoner's dilemma game, the chicken game, and the stag hunt game. In fact, our analysis included all 78 types of 2×2 (ordinal) games taxonomized by Rapoport, Guyer, and Gordon (1976). Similarly, for $n \times n$ games where the size of the game is greater than 2, we did not restrict our attention to a particular type of game. Payoffs for all games were generated by randomly sampling each payoff independently from a normal distribution (with a mean of 0 and a standard deviation of 100); therefore, every possible game has a nonzero probability of being played. We chose the normal distribution, rather than a uniform distribution, to capture the typically negative correlation between magnitudes in payoff (whether negative or positive) and the probability of their occurrence in the real world (Pleskac & Hertwig, 2014; see chapter 3).

5.4.2 The Properties of the Environment

An environment e is defined by the characteristics of the games that nature randomly (or exogenously) determines are to be played. Each environment is characterized by two variables: first, the size of the game, n (the players' action space), and second, the probability that each payoff in the game is missing or unknown to a player (in %). In an environment with m% missing observations, each individual payoff is randomly determined to be unavailable to the players with probability m. In other words, the player's state of knowledge is incomplete, inducing environmental uncertainty. This procedure creates a problem, however. When numerous payoffs are missing, it is unclear how to calculate best-response profiles. To solve this problem, one would have to hypothesize about how a player chooses an action under such circumstances. Because any hypothesis would (at this point) be quite arbitrary, we chose a different solution. Relying on the principle of vicarious functioning (see Brunswik, 1952; Dhami, Hertwig, & Hoffrage, 2004), players can use available cues (here, payoffs) to infer unreliable or unavailable cues (payoffs). For games with missing payoff information, we therefore assumed that the heuristic imputes a value equal to the mean of the player's other payoffs (with small perturbations to avoid complications associated with equal payoffs and ties). In evaluating the heuristics' performance, however, we assumed that the actual payoff a player receives, depending on the heuristic chosen, is based on the true payoff values and not on the thus inferred values. Let us emphasize that a Bayesian Nash equilibrium would require that players impute the whole probability distribution of possible values for the missing payoffs—in other words, they would need to calculate a decision rule repeatedly (a very large number of times) to approximate the distribution of each possible realization of the missing payoffs. Boundedly rational agents would be unlikely to carry out this calculation.

In sum, the set of environments E we investigated consisted of all possible combinations of the sizes of the action space and the likelihood of missing payoff information $\{n,m\}$, where $n \in N=\{2,3,\dots,19,20\}$ and $m \in M=\{0,5,10,\dots,70,75,80\%\}$. For example, a single environment $e(n, m)$ consists of a set of $n \times n$ games, sampled using the procedures noted above, with a probability m of payoff values being missing. Using this implementation, we could examine the performance of the heuristics as a function of the size of the game and the degree of missing knowledge (environmental uncertainty). For every environment, we simulated play for 10,000 randomly drawn games, thereby fully covering the game space. An extended analysis of a larger set of environments including different degrees of conflict or common interest between the players can be found in Spiliopoulos and Hertwig (2018).

5.4.3 The Competing Decision Policies

Our analysis compared 10 decision policies: a set of eight heuristics that have been identified as commonly employed by actual players in laboratory settings (Costa-Gomes, Crawford, & Broseta, 2001; Costa-Gomes & Weizsäcker, 2008; Devetag, Di Guida, & Polonio, 2016; Polonio & Coricelli, 2015; Spiliopoulos, Ortmann, & Zhang, 2018; Stahl & Wilson, 1994, 1995), the (normative) Nash equilibrium, and a baseline random policy. The decision policies are defined in table 5.2; more detailed examples of how to compute them will be presented shortly. To play games against simulated opponents using these decision policies, please visit interactive element 5.1 (at https://taming-uncertainty.mpib-berlin.mpg.de/).

We investigated the performance of level-*k* heuristics as well as somewhat naive heuristic solutions, such as choosing the action with the highest payoff. Level-*k* heuristics and cognitive hierarchy theory have been particularly successful in modeling boundedly rational behavior (e.g., Camerer, Ho, & Chong, 2004; Chong, Ho, & Camerer, 2016; Costa-Gomes et al., 2001; Nagel, 1995; Stahl, 1996; Stahl & Wilson, 1995). Each of the policies studied, with the exception of the (pure-strategy) Nash equilibrium, always suggests a unique action for any normal-form game (with unique payoffs). In order to resolve the coordination problem in the case of multiple Nash equilibria, we assumed that players choose the equilibrium maximizing the joint payoffs to both players. Some games may not have a pure-strategy Nash equilibrium;[1] we assumed that in such cases the Nash equilibrium strategy chooses an action randomly.

1. Although all finite-player, finite-action games—such as the ones we study in this chapter—are guaranteed to have at least one Nash equilibrium, for some games this will be an equilibrium in mixed strategies, not pure strategies. We focused solely on pure-strategy Nash equilibria because the calculation of a mixed-strategy Nash equilibrium is extremely computationally demanding.

Table 5.2

The set of simulated decision policies (excluding the random baseline): Eight heuristics and the Nash equilibrium.

Decision policy	Abbreviation	OwnP	EQW	NW	Dominance Self/	Dominance /Opp.	Description
Maximax	MaxMax	Y	–	Y	Y	–	Choose the action(s) offering the highest payoff for the player
Maximin	MaxMin	Y	–	Y	Y	–	Choose the action(s) offering the highest worst-case payoff for the player
Level-1	L1	Y	Y	–	Y	–	Choose the action(s) offering the best response to the assumption that an opponent is choosing randomly
Social maximum	SocMax	–	–	–	–	–	Choose the action(s) maximizing the sum of the player's own payoff and the opponent's payoff
Equality	Eq	–	–	–	–	–	Choose the action(s) minimizing the difference between the player's own payoff and the opponent's payoff
Dominance-1	D1	–	Y	–	Y	Y	Choose the action(s) offering the best response to the assumption that an opponent is choosing randomly over their nondominated actions
Level-2	L2	–	–	–	Y	Y	Choose the action(s) offering the best response to the assumption that an opponent is applying L1
Level-3	L3	–	–	–	Y	Y	Choose the action(s) offering the best response to the assumption that an opponent is applying L2
Nash equilibrium	NE	–	–	–	Y	Y	Choose the action(s) consistent with the Nash equilibrium

Note. OwnP: the heuristic considers only the player's own payoffs and ignores the opponent's payoffs; EQW: equal weighting of the opponent's payoff; NW: no weighting. Dominance Self/Opp.: obeys the principle of strategic self-/opponent-dominance (see box 5.1). Y: policy has the property; –: policy does not have the property.

The following examples illustrate how to compute each decision policy from the perspective of player A in the game presented in table 5.1.

Maximax (MaxMax). Identify the action offering you the highest possible payoff in the matrix (94). On this basis, choose action Up.

Maximin (MaxMin). Find the lowest payoff in each row—these are 32, 23, and 23 in the first, second, and third rows, respectively. Identify the maximum of these lowest payoffs (32). On this basis, choose action Up, thereby guaranteeing a payoff of at least 32.

Level-1 (L1). Sum your own possible payoffs in each row—these are 184, 94, and 134 for the first, second, and third rows, respectively. Identify the row with the largest sum (184). On this basis, choose action Up.

Social maximum (SocMax). For each cell, calculate the collective sum of payoffs to both players (e.g., 58 + 57 = 115 for the upper-left outcome). Identify the maximum total payoff (lower-left outcome, 144). On this basis, choose action Down.

Equality (Eq). For each cell, calculate the absolute difference in the players' payoffs (e.g., 94 − 41 = 53 for the upper-right outcome). Identify the minimum absolute difference (upper-left outcome, 1). On this basis, choose action Up.

Dominance-1 (D1). Examine whether your opponent has any dominated actions. Player B's Left action dominates their Right action: regardless of your actions, player B will always have a higher payoff from the former than the latter. Eliminate the dominated strategy (Right) from consideration and perform an L1 calculation on the remaining strategies (Left and Center). The possible payoffs are 58 + 32 = 90 for Up, 34 + 23 = 57 for Middle, and 70 + 41 = 111 for Down. Therefore, choose Down, which yields the highest mean payoff over player B's nondominated actions.

Level-2 (L2). Sum player B's payoffs for each action—these are 177, 132, and 110 for the first, second, and third columns, respectively. Assume that player B uses L1 and will choose the action with the highest sum, Left.

Identify your highest possible payoff in that column (70) and choose the corresponding action, Down.

Level-3 (L3). Assume that player B is an L2 player. By definition, an L2 player assumes that their opponent uses L1. Apply the L1 strategy to your own payoffs—as described above, this will result in the choice of action Up. As an L2 player, player B will respond by searching for their highest possible payoff in the row corresponding to Up, and will play Center. Choose the best response to this action by identifying your maximum possible payoff in the column associated with Center. The corresponding action is Down (41).

Nash equilibrium (NE). Identify the maximum payoff in each row (i.e., the best response to the assumption that player B has played each column). These are 70, 41, and 94 for actions Left, Center, and Right, respectively. The corresponding best response actions are Down, Down, and Up, respectively. Perform the same operations for player B's payoffs when you play each row. The corresponding best responses are Center, Left, and Left for your actions Up, Middle, and Down, respectively. Determine for which actions these two best responses coincide. In this case, this occurs for the combination Down, Left. Therefore, you would choose Down and player B would choose Left.

5.4.4 Classification of Heuristics' Paths to Simplification

Of the policies we simulated, the Nash equilibrium is the most complex computationally. The heuristics implemented represent different paths to reduce computational complexity: by reducing the amount of information required, by rendering the process of integrating information less complex, or by simplifying the assumptions made about an opponent or their beliefs. Spiliopoulos et al. (2018) measured the complexity of the eight heuristics examined in terms of the number of elementary information processing units required to execute each heuristic (see J. W. Payne, Bettmann, & Johnson, 1993; for an alternative approach to measuring complexity, using a cognitive architecture, see Fechner, Schooler, & Pachur, 2018). The paths taken to reduce complexity can be broadly categorized as payoff-based or probability-based simplification.

Payoff-based simplification. A heuristic can choose to completely ignore the opponent's payoffs and thus act as if the game were a nonstrategic task. The heuristics that do so are MaxMax, MaxMin, and L1 (see table 5.2; similar heuristics, applied to games against nature, are also discussed in chapter 8). All other heuristics in table 5.2 require information about both the player's own payoffs and the opponent's payoffs.

Probability-based simplification. Another path to simplification involves the beliefs held by a player about the probability with which the opponent will play each of the actions (see table 5.1). One simple assumption is that the probabilities of each action are equal, resulting in an equal weighting of the player's own payoffs—this can be viewed as a strategic variant of the equal-weighting principle proposed for prediction (Dawes, 1979). The heuristics that make this assumption are L1 and D1. The latter first removes the opponent's dominated actions, and only then assigns equal weights to the remaining actions. The L1 heuristic is of particular interest because it is known to be frequently recruited by real players in strategic games (see Costa-Gomes et al., 2001; Costa-Gomes & Weizsäcker, 2008; Devetag et al., 2016; Polonio & Coricelli, 2015; Spiliopoulos et al., 2018; Stahl & Wilson, 1994). It both ignores the opponent's payoff and assumes that the opponent randomizes over their actions with probability $1/n$. Alternatively, a heuristic can forgo beliefs about the opponent's behavior and focus solely on payoff information. This applies to SocMax and Eq. These two heuristics do not weight payoffs according to the likelihood of the relevant outcomes being obtained. Instead they involve either the addition or subtraction, respectively, of the player's own payoff and the opponent's payoff for every possible outcome, and then perform ordinal comparisons only. Finally, an even more Spartan process is not integrating the player's own payoffs with the opponent's payoffs but instead making a choice based on a single payoff. This approach avoids both weighting and adding and completely forgoes any belief formation. MaxMax and MaxMin belong to this computationally simplest class of heuristics. MaxMax chooses the action with the highest payoff for the player; MaxMin, the action with the highest payoff for the player in the worst-case scenario. Both heuristics involve only ordinal comparisons among the player's own payoffs based on max and min operations.

5.4.5 The Role of Strategic Dominance

The heuristics differ on another important dimension—namely, their adherence (or lack thereof) to the principle of *strategic dominance*—and their assumptions about whether or not the opponent adheres to this principle. If one of a player's (pure) strategies is better than another strategy—independent of the strategy chosen by the opponent—then that strategy *dominates* the other strategy (the latter is the *dominated* strategy). If a strategy dominates all other strategies in a game, it is referred to as a *dominant* strategy. We use the term *self-dominance* to describe the relationship between the player's strategies and *opponent-dominance* to describe the relationship between the opponent's strategies. Box 5.1 gives examples of the comparisons necessary to determine dominance in the game outlined in table 5.1. It is always beneficial to avoid a self-dominated strategy, regardless of any strategic uncertainty about the opponent's behavior, because a dominated strategy is, by definition, inferior to a dominant strategy for all possible actions available to the opponent. Of the heuristics in table 5.1, L1 and MaxMax adhere to the principle of self-dominance but do not assume that the opponent will do so. D1, L2, L3, and NE adhere to both strategic dominance principles. Whether the assumption of opponent-dominance is realistic depends on the heuristic employed by the opponent. It is not realistic if opponents use SocMax or Eq, neither of which systematically adheres to self-dominance.

Whether or not it is advantageous to adhere to the principle of dominance in a particular environment also depends on the probability that dominated actions exist in each of the environment's games. The proportion of games that have at least one dominated action quickly approaches zero as the size of the game increases from three onwards.[2] Therefore, the importance of recognizing dominance can be expected to decrease with larger games. For example, because L1 and D1 differ only in the first-step dominance check performed by D1, they will converge in their recommended actions as the number of actions increases (and the probability of a dominated action decreases).[3] However, in terms of processes and

2. The probability for $n=2, 3, 4, 5, 10$, and 20 is 0.5, 0.53, 0.48, 0.4, 0.08, and 0.0004, respectively.

3. Inspecting a game for the existence (or lack thereof) of dominated or dominant strategies can reveal further relationships between strategies. As observed by

information needs, D1 would still require more steps to arrive at the same recommendation than L1.

5.4.6 Measuring the Competing Decision Policies' Performance

We measured the success of each decision policy against a performance criterion that we refer to as the *Indifference criterion*. Before we describe this criterion, let us briefly outline the reasoning that informed its choice. In the simulations, we matched each of the competing policies (see table 5.1) as opponents in 10,000 randomly drawn games (for each environment) to approximate the performance of pairs of competitors within an environment. Each policy also played against itself. Imagine each player choosing in advance which decision policy to use across all games in a specific environment. Now consider how strategic uncertainty affects the choice of policy. If a player knows the opponent's policy, they can easily figure out the policy offering the best response. If they do not know the opponent's policy, they may form beliefs about the distribution of decision policies in the population and, at least in theory, calculate the expected payoffs for each policy conditional on those beliefs. This is a probabilistic quantity: it requires weighting the expected payoff against each decision policy by the probability of being matched with a player using that particular policy. However, as we argued before, it will be very difficult to learn this distribution. Consequently, players face strategic uncertainty—that is, they are unable to assign probabilities to the distribution of policies in the population. Luce and Raiffa (1957) argued that decisions in such large worlds (Savage, 1954) may be enabled by the principle of indifference (also known as the principle of insufficient reasoning). This principle informs our performance criterion.

According to the principle of indifference, each decision policy is equally likely to be used by the opponent (see also chapter 2 on the role of this assumption in individual choice). A player's expected payoff over the whole

Costa-Gomes et al. (2001), a level-k heuristic's proposed action is identical to that of the Nash equilibrium for games that are solvable by k rounds of iterated dominance (in pure or mixed strategies). Dk heuristics are identical to the Nash equilibrium in games that can be solved by $k+1$ rounds of pure strategy dominance. In many 2×2 games, there is significant overlap between strategies, enabling many simpler heuristics to emulate the Nash equilibrium.

set of games in an environment is then simply an average of the expected payoffs against each decision policy. We define this as the Indifference criterion.[4] It is the expected payoff a decision policy will achieve if (a) the policy plays against a population of policies that are uniformly distributed in the player population, or (b) it plays against a single decision policy, but is unaware which one it is, and believes that the policy is drawn with equal probability from the set of 10 decision policies.

More formally, let the action space A (A') of a player (opponent) denote the set of actions a_n (a'_n) available. The number of actions for each player (or size of the action space) is denoted by N and N'. We assume that this is the same for both players and therefore refer to the common value N as the size of a game. A normal-form game g is defined by a mapping from the action spaces of both players to payoff functions $\pi_g(a_n, a'_n)$ and $\pi'_g(a_n, a'_n)$. That is, the combination of actions (a_n, a'_n) determines the payoffs of both players. Let the decision policy space D (D') of a player (opponent) denote the set of policies d (d') available.

The first number in each cell of a normal-form game, as presented in table 5.1, represents player A's payoff; the second, player B's payoff. For example, if player A chooses a_1 and player B chooses a'_2, then the former receives a payoff of 32 and the latter 89. According to the Indifference criterion, denoted by $\bar{\pi}(d|e)$, the performance of a decision rule d in an environment e consisting of a set of games G_e is given by:

$$\bar{\pi}(d|e) = \frac{1}{|D'|}\sum_{d \in D'} E_{G_e}[\pi_g(a(d), a(d'))]. \tag{1}$$

As mentioned in section 5.4.5, the L1 and D1 policies make identical choices for games without dominated strategies. Consequently, the L2 policy is the best response to both of these policies in such games, giving it an unfair advantage over other policies (which at most are best responses to only one other policy; e.g., NE is the best response to NE, L3 to L2, L1 to Random). We leveled the playing field by assigning half the weight to L1 and D1 when calculating the Indifference criterion. Including both with

4. A larger set of performance criteria, including one based on Wald's (1945) maximin model and another examining the robustness of policies to different (nonuniform) distributions of policies in the player population, can be found in Spiliopoulos and Hertwig (2018).

full weight does not alter our findings significantly, although as expected it does give L2 a small boost in performance.

In a nutshell, the Indifference criterion captures policies' average performance against the whole set of possible opponent policies, assuming each is equally likely. We now turn to the results of the competition to examine how well or poorly the heuristics fared when facing both strategic and environmental uncertainty.

5.5 Decision Policies' Performance as Measured by the Indifference Criterion

Figure 5.1 shows the performance of the competing policies across the set of environments, measured in terms of the Indifference criterion. The size of the action space (or game) n is shown along the y-axis; the percentage of missing information, along the x-axis. Each heatmap depicts the performance of one policy. As a benchmark, the two top-left panels show the performance of random choice and the Nash equilibrium. The darker the shading, the better the performance. The robustness of a policy is a function of how large the area of dark shading is: the larger the area, the more robust the policy's performance over a range of game sizes and degree of payoff uncertainty. In general, the findings presented in figure 5.1 suggest that some heuristics exhibit consistently good performance over the whole set of environments and can even outperform the Nash equilibrium. Other heuristics perform well only in specific regions. We now move on to summarize the key results. Recall that performance when operationalized in terms of the Indifference criterion (figure 5.1) assumes that each opponent policy is equally likely. The effects of changing the distribution of decision policies in the population can be viewed in interactive element 5.2.

The first result is that assuming equal weighting pays off. The L1 and D1 heuristics, both of which assume equal weighting of the player's own payoffs, are the best performers on average across all the environments, and show robustness to limited payoff knowledge regardless of the size of the game. The next best policies (averaged across all the environments) are the SocMax and L2 heuristics. Note that the former outperforms the L1 and D1 heuristics in niche environments comprised of large games and high payoff uncertainty.

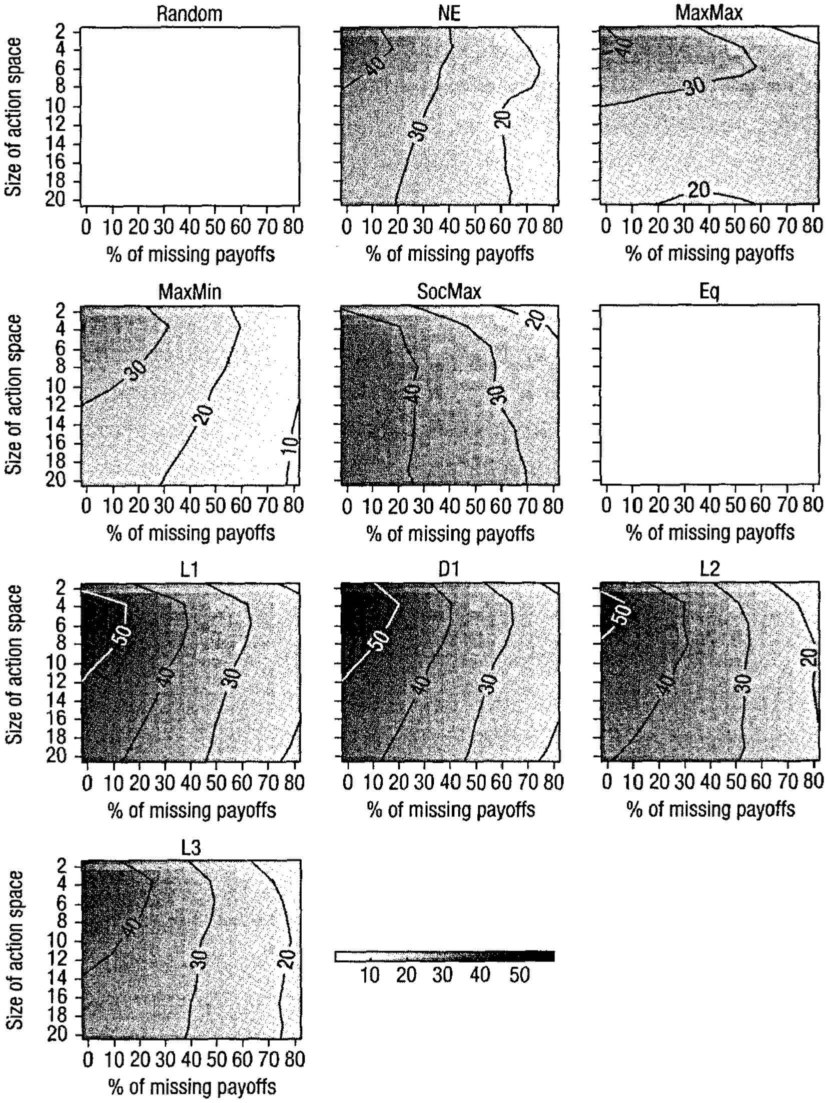

Figure 5.1

Results of the competition among policies, with performance operationalized in terms of the Indifference criterion. The average percentage of missing payoffs ranges from 0% to 80%; the size of the action space ranges from 2 to 20. The darker the shading, the better the performance.

The second result is that there is such a thing as too much simplification. Two of the computationally simplest heuristics, MaxMax and MaxMin, do not perform well. They achieve moderately high mean payoffs in just a small subset of environments, namely, games involving few actions (2–5) and low payoff uncertainty. The extreme simplification embodied in these heuristics seems to overshoot the mark, leading to poor performance.

The third result is that assuming higher rationality in the other player leads to poorer performance. The L1 and D1 heuristics are the most robust policies, achieving excellent performance across the environments. Their performance decreases only when knowledge about payoffs is extremely limited—however, all policies' performance declines under this condition. Increasing the level of rationality attributed to the opponent leads to a further decrease in performance. Although the L2 and L3 decision rules still perform reasonably well, they are less robust than L1 and D1, especially as game size increases. Similarly, the Nash equilibrium policy, which makes the strongest assumptions about the opponent's rationality, performs worse than the L2, L3, and L1 and D1 heuristics (which attribute a lower level of rationality to the opponent). Importantly, the Nash equilibrium achieves relatively high payoffs only in very small games with low payoff uncertainty.

The final result is that the top-performing policies obey the self-dominance principle. According to the Indifference criterion, the best performing heuristics are L1 and D1, in terms of robustness to both limited knowledge and game size. Both heuristics obey the self-dominance principle; D1 also assumes the opponent obeys the dominance principle. Note that the L2 heuristic, which also performs well, obeys both self- and opponent-dominance.

5.6 The Robust Beauty of Simplicity

We investigated the performance of a variety of policies in the face of strategic and environmental uncertainty. All alternatives to the normative Nash equilibrium policy simplify the decision process. Some, such as L1 and MaxMax, make strong simplifying assumptions and, for instance, take no account of the opponent's strategic concerns (see table 5.1); others, such as D1 and L2, take those concerns into account, albeit in a somewhat simplified manner, by making weaker assumptions about an opponent's rationality or common knowledge of rationality. We also investigated the heuristics' performance as a function of environmental uncertainty, that is, the degree of knowledge about the player's own payoffs and the opponent's payoffs. Finally, we varied the games' complexity, that is, the number of actions available to each player.

How did the heuristics fare? Consistent with findings on inferential as well as preferential choice (Gigerenzer, Hertwig, & Pachur, 2011; see chapter 2), simplicity was more robust to strategic and environmental uncertainty than was complexity. The normative Nash equilibrium policy was only relatively competitive for precisely the kind of games typically used in experimental investigations of strategic behavior: complete knowledge of payoff information and game sizes of roughly 2–5 actions. Furthermore, making specific assumptions about an opponent's behavior (e.g., predicting, as an L3 player, that the opponent will behave as an L2 player) may risk misrepresenting actual heterogeneity among players. By contrast, the equal-weighting L1 and D1 heuristics, both of which attach the same likelihood to each of an opponent's possible actions (the latter after removing dominated actions), were the best performers, followed closely by L2 and SocMax.

The L1 heuristic is particularly adapted to environments in which knowledge about payoffs is severely limited. It performs very well, despite turning a strategic task into a nonstrategic one by completely ignoring the opponent's payoffs. Echoing results in games against nature, there is also a "robust beauty" (Dawes, 1979, p. 571) to this improper decision policy in strategic games.

5.7 Sterelny's Error

Sterelny (2003) argued that simple rules of thumb are likely to fail in competitive interactions. Yet we found that even in environments marked by both strategic and environmental uncertainty, heuristics that abandon normative axioms and Bayesian principles need not buckle. In fact, it was typically sufficient that a heuristic obeyed the self-dominance principle, which even the simple L1 heuristic does, and refrained from making precise assumptions about an opponent's behavior. Two simplifications were particularly successful: equal weighting and ignoring an opponent's payoff (or, equivalently, the strategic component of the game). But let us also emphasize the limits of oversimplification. Decision policies that base decisions on a single piece of information without weighting, such as MaxMax and

MaxMin, typically performed worse than other heuristics in terms of the Indifference criterion. Apart from oversimplification, being overly confident in the ability to predict how an opponent is likely to play also represents a risk. For instance, precise assumptions or beliefs about what other players will do often came at a high price. The L3 heuristic and the Nash equilibrium are cases in point.

If anything, our analysis may have underestimated the performance of simple heuristics relative to more complex processes in strategic interactions. First, we assumed that each policy was perfectly executed. Yet the more complex a policy is, the more difficult it will be for it to achieve perfect execution. Consequently, the Nash equilibrium and the relatively complex high level-*k* heuristics are more susceptible to execution errors. If these errors were random, the performance of policies prone to execution errors would be a linear combination of their performance assuming perfect execution and the performance of the random choice rule (worst competitor). Thus, execution errors would further attenuate the performance of more complex policies relative to simpler ones such as L1. Second, a policy's complexity is likely to be positively correlated with its execution time. Therefore, a decision maker using a heuristic would have more time available to play more games. In other words, if performance were normalized per execution time, the performance of simpler policies would increase relative to that of more complex policies.

5.8 Conclusion

Experimental studies of strategic interactions have found evidence for the frequent use of heuristics (see Spiliopoulos & Hertwig, 2018). Our results cast new light on this finding. The established narrative attributes it to humans' inability to reason according to the complex Nash equilibrium. Alternatively, it has been argued that people still optimize their choice of policy but do so subject to a constraint based on the decision costs, leading to the use of simpler policies. This argument is based on the accuracy–effort (speed) trade-off, often seen as a general law of cognition (see chapter 2): those who invest less mental effort will pay a price in terms of lower accuracy (performance). From this perspective, heuristics are, by definition, always second best; people use them because they take a rational approach to their cognitive limitations. Were resources unlimited, more computation and more time would, from this perspective, always be better. Our analyses show that the accuracy–effort trade-off is not ubiquitous. Simpler solutions such as level-*k* heuristics—in particular, L1 and D1—do not inevitably sacrifice performance relative to the Nash equilibrium policy. In fact, they can achieve high performance and robustness in the face of substantial environmental and strategic uncertainty. This does not mean, however, that there are no limits to the benefits of simplicity. As we have also observed, simplifying assumptions can be too naive and too minimal.

It has often been argued that "choice in social interaction harbors a level of complexity that makes it unique among natural decision-making problems, because outcome probabilities depend on the unobservable internal state of the other individual" (Seymour & Dolan, 2008, p. 667). We agree that social environments are different from physical ones to the extent that the presence of others along with their strategic intentions and counterstrategies represent additional and important sources of uncertainty. Yet our results challenge the common wisdom that social complexity necessitates cognitive complexity (e.g., Humphrey, 1976, 1988; Whiten & Byrne, 1988). Even in environments fraught with environmental and strategic uncertainty, simple heuristics can reach surprisingly accurate decisions and prevent people from making overly bold, demanding, and specific assumptions about others.