The Use and Abuse of Game Theory in International Relations

THE THEORY OF MOVES

RANDALL W. STONE

Department of Political Science
University of Rochester

The author argues that the theory of moves, which has gained popularity in recent years as an alternative to game-theoretic analysis of strategic interaction, is fundamentally flawed. The theory’s adherents argue that it makes theoretical progress by endogenizing the structure of games and introducing new ways of analyzing repeated interactions. The author analyzes the theory of moves from a game-theoretic perspective and challenges its theoretical claims. The author then reanalyzes several recent articles that have used the theory of moves, showing that its application to empirical cases is strained and that game theory can provide models that do a better job of fitting the stories the authors tell about them.

The theory of moves seeks to offer a new way of analyzing the $2 \times 2$ games that have become a ubiquitous feature of how we think about international relations. In particular, they hope to inject a dynamic element into game theory, which Steven Brams (1994) and his collaborators regard as an essentially static enterprise (Zeager and Bascom 1996, 471). Brams believes that he has developed a powerful new tool for reanalyzing the strategic situations that international relations specialists have considered important in our field and that his approach is able to capture nuances of politics that game theory cannot. Brams claims that (1) the theory of moves generates new insights about the familiar $2 \times 2$ games commonly used in international relations; (2) the theory of moves captures nuances of strategic interaction that game theory misses because game theory treats the extensive form—who gets to move when—as exogenously determined, whereas the theory of moves makes it endogenous; and (3) the theory of moves makes theoretical progress in understanding repeated games by exploring dimensions of power—"threat power" and "moving power"—that game theory cannot grasp (Brams 1994, 6-7, 17-18).

I will show below that each of these claims is mistaken. The problem, however, is not that the authors of the theory of moves fail to understand game theory—indeed, some of them have contributed sophisticated game-theoretic analyses to the international relations literature—but that they have misunderstood what happens when they
attempt to apply the theory of moves. In the first section, I provide an example of Brams’s method by discussing his analysis of the familiar game of chicken. This example illustrates the point that the theory of moves does not, in fact, analyze the situations familiar to international relations specialists as $2 \times 2$ games; instead, it substitutes sequential games for the simultaneous ones. In addition, I show that the theory of moves does not, in fact, make the structure of these games endogenous, as its adherents claim. In the second section, I discuss Brams’s attempt to introduce a dynamic element into his method by using the concept “moving power,” and I analyze his treatment of the Arab-Israeli conflict in these terms. I argue that moving power is a loose way to describe the dynamics of repeated games and, when applied as an explanation, is tautological. In the third section, I discuss the use of Brams’s method in recent articles by other authors and show how the cases that they describe can be reanalyzed in game-theoretic terms.

I conclude that the theory of moves represents backsliding rather than progress in the effort to understand strategic interaction. At worst, using the theory of moves means using the wrong model. Because it replaces a simultaneous-move game with a sequential game, the theory of moves cannot analyze any situation involving strategic uncertainty. Because the sequential game that it generates has an arbitrarily assigned structure, there is no guarantee that the model generated captures the strategic features of the situation that is being analyzed. The theory of moves is at its best when the strategic situation to be analyzed really is sequential and the algorithms specified by the theory of moves generate a plausible structure for the game. Even then, however, the cost of employing the theory of moves is that we are forced to renounce the game-theoretic tools that have allowed us to reason with precision about strategic interaction and derive conclusions about issues such as reputation, information, and credibility.

### THE THEORY OF MOVES IN ACTION: THE CASE OF CHICKEN

In game theory, a game matrix represents a strategic situation in terms of choices that must be made simultaneously. This is referred to as the normal form of the game. To analyze sequential choices, game theory ordinarily uses extensive-form games (i.e., game “trees”). However, it is possible to convert any extensive-form game into a normal-form game that preserves the structure of the extensive form in which rows and columns represent plans for how to play the extensive-form game from beginning to end. The theory of moves departs from these conventions. It treats a game matrix as a kind of map on which cells represent stages in a sequential game, and players can move around by changing their strategies. The payoffs in the cells are hypothetical, so there is no accumulation of payoffs during play; it is only the payoff in the final cell that matters. Practitioners of the theory of moves first choose an arbitrary “starting point” for the game—one of the four cells of a game matrix—and an arbitrary “first mover.” From this initial point, the two players move around the game matrix like rooks chasing each other around a chess board, until one voluntarily stops moving.
As an illustration of Brams’ method, I compare the game-theoretic treatment of a game familiar to international relations scholars with the treatment by the theory of moves. The game known as “chicken” is commonly used to analyze crisis bargaining. The story behind the game is that two socially maladjusted youths of my parents’ generation are barreling toward each other in jalopies in a misguided attempt to demonstrate their courage. At the last possible instant, each is faced with a terrible dilemma: if neither swerves, both will surely be killed. On the other hand, whoever swerves will be branded a “chicken” and will be ostracized and humiliated, whereas the other will gain in social status. If both swerve, the outcome will be less dire, but neither will enjoy the benefits of prevailing in the dispute. Game theory depicts this situation in Figure 1.

CHICKEN ACCORDING TO GAME THEORY

The two forms of the game depicted in Figure 1 are equivalent. The game matrix presents the game in normal form; the outcome is determined by the simultaneous choice of C (swerve) or D (don’t swerve) by each player, yielding the payoffs shown in the box where their strategies intersect. The payoffs to row are listed first, so if row plays D and column plays C, row gets its best payoff (4), and column gets its second-worst payoff (2). The game tree presents the same strategic situation in extensive form: paths leading to the left indicate the choice to swerve, and paths leading to the right represent the choice to barrel ahead. The dotted line connecting the two nodes where column chooses (called an information set) indicates that column does not know what decision row has made or, equivalently, that the choices are made simultaneously.

The key feature of this game that makes it analytically interesting is that the choices have to be made simultaneously—literally at the last possible instant. There is no dominant strategy for either player (no best way to play that does not depend on the other’s choice) because each would prefer to swerve if the other does not and not swerve if the other does. The game has three Nash equilibria. In pure strategies, the pairs (C,D) and (D,C) are both equilibria because if either player is convinced that the other will not swerve, he or she surely will, and if either player is convinced that the other will swerve, he or she certainly will not. There also exists a mixed-strategy equilibrium in which each player chooses to swerve with a certain (commonly known) probability, which is exactly chosen to make the other indifferent between swerving and not swerving.¹ The mixed-strategy equilibrium may seem contrived to some readers, but it is intuitively the most compelling because it is hard to see why anyone would choose to play such a game with the intention of swerving, and it captures the empirical reality that chicken is a very dangerous game that often results in casualties.² As a model of superpower crises, the mixed-strategy equilibrium captures Schelling’s (1960) notion of a strategy that “leaves something to chance,” which, he argues, was essential to make extended nuclear deterrence credible.

¹. Note that the precise probabilities chosen depend on the cardinal payoffs because we need to calculate expected utilities to determine the optimal strategies. However, such an equilibrium will exist for any set of cardinal payoffs consistent with the ordinal rankings of chicken.

². Neither of the pure-strategy equilibria is characterized by any risk because a disaster cannot occur. In the mixed-strategy equilibrium, a disaster occurs whenever both players randomly choose to barrel ahead.
Now, suppose that we take a different model that has the same preference orderings but in which row moves first, observed by column, and then column moves. I will call this game “sequential chicken” (Figure 2). Note that the only difference between this game and chicken (Figure 1) is that there is no dashed line, which indicates that in this model, column knows row’s move before choosing its own. Although this game looks similar to the game of chicken analyzed above, the similarity is superficial; the strategic situation is completely changed by the introduction of sequential play. This is the claim that Schelling (1960) makes in his discussion of chicken and credible commit-
ments: he asks what would happen if one player were able to precommit to playing D—for example, by dramatically hurling the steering wheel out of the driver’s window—and concludes that this would force the other player to concede. Consistent with Schelling’s intuition, once column observes that row has chosen D, there is a sense in which column no longer has any choice but to play C (otherwise, a disaster ensues). Row anticipates that this will be the case and consequently has no reason to choose anything but D. Formally, in this game, (D,C) is the unique subgame-perfect equilibrium (SPE).

Showing how game theory describes the normal form of sequential chicken will help to demonstrate how it differs from chicken. By definition, the normal form and the extensive form of a game must depict the same strategic situation. Consequently, the normal form of sequential chicken must be different from the normal form of chicken and incorporate the fact that actions in sequential chicken are sequential. Note that sequential play provides column with two new strategic options: as before, it can choose categorical strategies, “always C” or “always D”, which do not take its observation of row’s move into account; or it can choose contingent strategies, “play the same choice as row” or “play a choice different from row’s.” The normal form of sequential chicken, consequently, is a $2 \times 4$ matrix (Figure 3).

Examining this normal-form representation of the game makes it clear that (D,C) is not the only Nash equilibrium. There are three Nash equilibria of this game (underscored). Note that a Nash equilibrium exists if both players’ strategies are optimal given the other’s expected strategy—regardless of whether the other’s strategy is cred-
There are two ways to get the (D,C) outcome: either column plays “always C” and row exploits this, or column plays “different,” and row exploits this. Alternatively, column can choose “always D,” which compels row to choose C. The problem with this equilibrium is that it relies on column’s ability to make an incredible threat: if row chose D, column would no longer wish to carry out its strategy and usher in a disaster. That is why the (C,D) equilibrium is not subgame perfect. Nevertheless, this is a Nash equilibrium because if row, for some reason, believed that column would carry out its threat, this would be the only sensible way for row and column to play.

Brams’s (1994) analysis is completely different because the theory of moves substitutes a sequential model of choice for the simultaneous model of chicken. As we have just seen, substituting a sequential model for a simultaneous model is a dramatic change: the model no longer represents the original strategic situation. Consequently, it is no surprise that Brams’s analysis of the same configuration of preferences leads to very different results.

### CHICKEN ACCORDING TO THE THEORY OF MOVES

Suppose that play begins at the upper-right-hand corner of Figure 1 (2,4), and row has the first move. According to the theory of moves, row can accept (2,4) or change its strategy to D, leading to a disaster (1,1); the disaster is painless, however, because column will surely shift its strategy to C to avert it, leading to (4,2). Row is satisfied by receiving its best possible outcome and stops play at this point. By moving first, row has forced its opponent to capitulate. Figure 4 translates the series of choices generated by the theory of moves into the more familiar form of an extensive-form game. Branches leading to the left represent a choice of C in the theory of moves, and branches leading to the right represent a choice of D. Row either accepts column’s preferred outcome or continues; if row continues, column is faced with a fait accompli and must choose between row’s preferred outcome and disaster. The equilibrium path is

---

3. Subgame perfection is a higher standard, which requires that every strategy be a Nash equilibrium starting from any point in the tree, including those that may never be reached. In other words, it asks the question, If your opponent didn’t act as expected, would it still be rational to carry out the strategy that you have threatened? In general, to find all of the Nash equilibria of a game, convert it to normal form; to find only the subgame perfect equilibria, convert the game to extensive form and solve it by backward induction.
Using backward induction, column chooses to capitulate, so row chooses to continue.

On the other hand, suppose that column has the first move. Column anticipates that it can be driven to (4,2), but by moving away from its own best outcome, it can preempt row’s stratagem. When column shifts to C, securing (3,3), it forces row to stop. If row chose to continue by moving to D (2,4), column could shift back to D, forcing a disaster (1,1). Again, however, the disaster would be painless because row could capitulate (play C), and column would return to its favored outcome (2,4). Play would terminate at this point because it had returned to the point of origin, and row would be unable to manipulate the agenda further. Since row prefers (3,3) to (2,4), row stops after column’s first move. It is more intuitive to represent this series of choices as an extensive-form game (Figure 5). Again, in terms of the strategies in the theory of moves, each branch leading to the left represents a choice of C, and each branch leading to the right represents a choice of D. The equilibrium path is bold. The right-hand branch of the tree is the same as Figure 4: if column does not move, row does, with the consequence that column is forced to capitulate. Column wishes to avoid this branch of the tree. The subgame that begins with column’s second move on the left-hand side is a
mirror image of Figure 4: if play reaches this point, column can force row to capitulate. Row wishes to avoid this outcome. Consequently, if column chooses the left branch of the tree, row will choose to stop at (3,3) rather than continue and capitulate at (2,4).

Notice three things. First, the structure of the game changes dramatically depending on the starting point and who moves first. Second, depending on the starting point and who moves first, (3,3) can often be supported as a “non-myopic equilibrium” (Brams 1994), although it cannot be supported as a Nash equilibrium of the normal-form game, chicken. Third, the logic behind the analysis is driven by the geometry of the game matrix and not by the logic of the strategic situation of which chicken is a model.

The structure of the game changes depending on the starting point and who moves first. One might conclude (prematurely, as we shall see) that the lesson that Brams’s (1994) method draws from chicken is that the player who gets to move first has the advantage. Indeed, this would be an intuitive lesson to draw from the chicken preference ordering, as we saw from the example of sequential chicken discussed above. This was Schelling’s (1960) point about the strategic value of being able to precommit to a policy of confrontation. However, this is not a general lesson that the theory of
moves can draw from chicken because this logic depends on the game beginning at the (4,2) or (2,4) positions. If play begins instead at (1,1), the first mover is disadvantaged because it is the first mover who will be faced with the dreadful choice of capitulation or disaster. If play begins at (3,3), neither player has an incentive to depart since that would allow the other player to barrel ahead, forcing the first to choose between capitulation and disaster. In the case of chicken, therefore, the theory of moves cannot conclude generally whether it is better to move first or second.

Indeed, the theory of moves does not generally find a characteristic logic to particular configurations of preferences because it generates from any given configuration numerous alternative extensive-form games that may have quite different properties. Because everything depends on the arbitrary assignment of a starting point, the theory of moves cannot say anything that is generally true even about a simple game like chicken.

(3,3) can be supported as a nonmyopic equilibrium. In fact, the theory of moves can support any Pareto-optimal outcome. In the case of chicken, the theory of moves finds that whenever play begins or arrives at (3,3), play stops, and (3,3) becomes the outcome. However, (3,3) is supported by a set of off-the-equilibrium-path strategies that seem very odd in terms of the original strategic situation of which chicken is a model. In terms of the theory of moves, if the row player defects, leading to its first-best state of the world, column can counter by defecting, which leads to disaster. In game theory, this is discounted as an incredible threat since no one prefers disaster to capitulation. In the theory of moves, however, the threat is costless because the outcome does not stay at disaster. Instead, the row player is forced to choose between disaster and capitulation and chooses to capitulate. Since row anticipates that this sequence of events leads to an outcome that is worse than the starting point and column does the same, no one departs from the initial state of the game.

If an equilibrium is a logical argument about how rational players ought to play a game, what is the substantive interpretation of this argument? Here, it is the claim that moving first makes one vulnerable to an ultimatum with which one will certainly comply, so no one will ever choose to move. Written as an extensive-form game, the series of choices spelled out by the theory of moves is illustrated in Figure 6. Figure 6 is the left-hand branch of Figure 5. Row chooses to stay at (3,3) or continue; column chooses to stay at (4,2) or continue; row chooses (2,4) or (1,1), and the game ends. The equilibrium path is in bold. By backward induction, column knows that row will choose (2,4) over (1,1), so column continues. Knowing that column will continue if given the chance and play will end at (2,4), row chooses to stay at (3,3).

In short, what we have here is a new game. This is not a reasonable hypothesis about how to play in a situation of simultaneous choice; instead, it is an entirely different strategic situation. It might represent, for example, how one nuclear power thinks about launching a conventional attack against another. If it launches the attack and the other does not respond with nuclear strikes against its military forces, it will gain an advantage; however, if the defender responds with tactical nuclear weapons, the attacker will face the unpalatable choice of retreating with heavy losses or initiating a full-scale thermonuclear war. Given this choice, the potential attacker would back
down, so it is deterred from launching the attack in the first place. This is not a bad model, and it might represent the logic of particular crises better than chicken. However, it is not a model in which simultaneous choices are made, so it cannot represent strategic situations in which simultaneous choices play an important role.

Compare this to the logic by which game theory arrives at a cooperative outcome in repeated chicken. There are numerous ways to achieve mutual cooperation in repeated chicken, and one of these is quite similar to the logic traced out by the theory of moves. In this equilibrium, each player, in effect, promises the other to play cooperatively until the other deviates; if one player deviates, the other promises to deviate forever. Given that I expect you to play this way, I know that I can get the high payoff from unilateral defection exactly once, but thereafter I would be forced to choose between a series of disasters and a series of capitulations, and I know that I would capitulate. Provided that I value the future sufficiently, I will prefer to cooperate every time rather than cheat.
once and pay for it for eternity, and the same logic applies to you. This can be shown to
be an SPE. 4

This particular repeated-game equilibrium relies on a threat similar to the one the
theory of moves uses in the extensive-form game represented in Figure 6. Is the theory
of moves, then, in effect analyzing the repeated game? The answer is no, for two rea-
sons. First, this is only one of many repeated-game equilibria. Others include repeating
either of the pure-strategy equilibria, the mixed-strategy equilibrium, or any combina-
tion of those and mutual cooperation. Where the theory of moves offers a unique solu-
tion, game theory insists that there are no unique solutions in repeated games. This is
not merely a limitation of game theory; it is a real-world problem, which the theory of
moves conveniently ignores. Second, game theory makes the important caveat that the
equilibrium is only sustainable if the players value the future sufficiently. The theory
of moves, on the other hand, cannot say anything about the credibility of repeated-
game equilibria.

The logic of the analysis is not related to the logic of chicken. The original game of
chicken was interesting not because of its geometry but because it represented a strate-
gic situation that was believed to be empirically significant. It is a model of simulta-
neous decision making in which there is a conflict of interest and the potential for
mutual disaster. None of the extensive-form games that the theory of moves proposes
to substitute for chicken, however, represents this situation. Instead of a dangerous
game in which both players move simultaneously and try to guess what the other will
do, the theory of moves suggests a menu of sequential games, each of which has
exactly one possible outcome. At this point, it is useful to return to the first two claims
that Brams (1994) makes for his method.

Claim 1: The theory of moves generates new insights about the familiar 2 × 2 games
commonly used in international relations. Ironically, as we have seen, one class of
games that the theory of moves can tell us nothing about is the one that it purports to
take as its subject: the familiar 2 × 2 games. In game theory, a 2 × 2 matrix represents a
simple normal-form game: the strategies are chosen simultaneously and independ-
ently. The theory of moves, however, cannot model situations involving simultaneous
choices—or, equivalently, choices that have to be made without knowledge of the
other player’s previous choice—because it converts every situation into a sequential
game of perfect information. Consider the example of the most famous of 2 × 2
games, prisoner’s dilemma: the district attorney separates two suspects and offers each
the same plea bargain. This is a device that has broken the most hardened criminals,
and the whole trick is to force them to make their choices independently.

Brams (1994) claims that situations involving simultaneous choice are rare in inter-
national relations, so little is lost. In point of fact, a basic game-theoretic insight is lost:
sometimes, decisions do have to be made under strategic uncertainty, and the results

4. An anonymous reviewer asked whether this equilibrium is renegotiation proof. It is because
the reversion strategy to punish a defection is to always defect. Because this strategy profile obliges the oppo-
nent to cooperate, it assures the punisher of his or her best possible outcome. Once you have triggered this
series of punishments, there is nothing that you can offer me that will make it worthwhile for me to relent.
are often tragic. In the theory of moves, there are no tragedies since everyone has an incentive to avoid them. In fact, the rules of the theory of moves assure that any Pareto-optimal outcome in a matrix can be sustained as a nonmyopic equilibrium from some starting point because someone will always be disadvantaged by moving away from (or failing to move to) a Pareto-optimal outcome. Romeo and Juliet live happily ever after. Pareto-inefficient outcomes, however, are the basic stuff of international relations: wars, trade disputes, alliance tensions, defaults, revolutions, terrorism. These are all inefficient outcomes. A theory that fails to predict inefficient outcomes appears to be rather irrelevant to our core concerns.

Brams (1994, 17-18) claims that rewriting his models in game-theoretic terms misses his point: he intends to discover patterns of behavior that are correlated with “game configurations” rather than with particular games. By game configuration, he means a particular $2 \times 2$ matrix, which generates numerous extensive forms—eight, for example, using the rule that play stops if it returns to the starting point. In other words, he wants to draw conclusions based on preference orderings alone that would be valid irrespective of the extensive form of a game. In general, however, this is impossible. The extensive form really does matter. The theory of moves generates arbitrary extensive-form games from any game matrix, and these differ depending on the starting point and the identity of the first mover. Because these extensive forms may be quite different from each other, there is no reason to expect the theory of moves to find a characteristic logic to particular preference configurations; indeed, it generally does not. In the case of chicken, we saw that the strategic situations that the theory of moves generated differed dramatically depending on the starting point and who moved first and that the theory of moves cannot reach any general conclusions about the strategic implications of this set of preferences.

Claim 2: The theory of moves captures nuances of strategic interaction that game theory misses because game theory treats the extensive form as exogenously determined, and the theory of moves makes it endogenous. It is true that game theory makes the extensive form—the order in which events may occur in a game—exogenous; in fact, one might say that the objective of game theory is to determine how the structure of games, together with the preferences and beliefs of the actors, determines strategic interaction. This does not mean, however, that decisions about “who moves first” cannot be left up to the actors. The model can specify that one player has a choice to move first or second or can introduce a random device, vote, or any other mechanism one wishes to propose. Game theory simply insists that these modeling choices have important consequences, so they must be made explicit. It is a curious misconception, however, that the extensive form is somehow made endogenous in the theory of moves. In fact, by selecting a player to make the first move, a starting point on the matrix, and a “stopping rule” to prevent endless cycling, the practitioner of the theory of moves

5. This does not stop Brams (1994) from making generalizations about game configurations: for example, some configurations allow cycling in the theory of moves, and some do not. However, all of these distinctions operate only within the rules of the theory of moves. It is not the case, for example, that all games generated from a particular matrix have a similar logical structure or could be used to represent similar strategic situations.
generates a unique extensive form for every game. The sequence of choices available to the two players can be written down as an extensive-form game, and the solution concept used in the theory of moves is familiar to game theorists as backward induction. The structure of the game is not endogenous; it is simply arbitrary. Changing the starting point, the player who makes the first move, or the stopping rule will generate a different exogenously fixed extensive form.\(^6\)

The implication is that the theory of moves cannot generate any insights that cannot be replicated by game theory since the repertoire of strategic models available in the theory of moves is a subset of the set of extensive-form games. Any strategic situation that can be modeled with the theory of moves can also be modeled as an extensive-form game. The converse is not true, however: the theory of moves can discuss only a small portion of the infinite variety of strategic situations that can be described by game theory. Moreover, the advantage of modeling situations as extensive-form games is that the structure of the strategic interaction directly determines the structure of the game; under the theory of moves, by contrast, the extensive form is assigned arbitrarily.

**THE SEARCH FOR A DYNAMIC ELEMENT: CYCLING IN ASYMMETRIC GAMES**

Brams (1994) seeks to extend his theory of moves to cover dynamic games—games with a time element in which players make a series of choices—by relaxing his assumption that play stops whenever an initial "state" is revisited. If the actors can continue moving, however, it is not obvious when they should stop "cycling" around the matrix. To impose some order, Brams assumes that a player always stops if he or she happens to be receiving the best possible payoff when it is his or her turn to move. This prevents cycling in all symmetric games (games in which the players’ ordinal payoffs are mirror images of each other, such as prisoner’s dilemma or chicken). In a number of asymmetric games, however, it is possible for cycling to occur. To achieve determinate predictions in these cases, Brams creates new rules, which he calls threat power and moving power.

When threat power comes into play, the players act as if the game were about to be infinitely repeated. The player with the power to make a threat is committed to playing a particular strategy, even if it would not be Nash-equilibrium behavior in the single-shot version of the game. For example, in his discussion of prisoner’s dilemma, Brams says that a player with threat power could commit to playing cooperatively in the first round and then to retaliating in hypothetical future rounds if the second player failed to reciprocate. As long as both players act as if the game will be repeated, the threat is effective, and both players cooperate (Brams 1994, 139). This is by no means a new

\(^6\) Brams (1994, 53-57) seeks to overcome this arbitrariness by inserting an “anticipation game” prior to play of the actual game and finds that, in many cases, play of the prior game allows him to choose a starting point in the one that follows. This, of course, does not make the extensive form endogenous, either; it simply substitutes an arbitrary choice of a prior game for an arbitrary choice in the following one.
insight. In developing his concept of threat power, Brams relies on the established game-theoretic result that a wide range of cooperative equilibria can be sustained in repeated games despite being impossible to reach in the single-shot version of the same games. The difference in approach between game theory and the theory of moves could not be sharper, however: whereas game theory rigorously proves that such equilibria exist and finds a precise set of strategies, beliefs, and necessary conditions on the players’ preferences to sustain them, the theory of moves simply assumes away the complex problems of reputation and credibility. In Brams’s (1994, 140, emphasis added) words, threat power “means that there is always later play that allows a threatener to recoup losses it may have incurred earlier in carrying out threats. Carried-out threats, I assume, enhance a threatener’s credibility, enabling it to deter future challengers.” In game-theoretic terms, threats involving threat power are subgame-perfect strategies by definition.

Moving power, according to Brams (1994), allows one of the players to continue to cycle through alternative states longer than the other, compelling the player without moving power to eventually stop moving on one of the two “states” at which it has the next option to move. This definition is rather loose because Brams does not explicitly model the passage of time and does not allow payoffs to accumulate during the cycling process. Instead, he makes an informal analogy to game-theoretic bargaining models in which delay is costly and the player with the greater resolve is able to gain a larger share of the prize (Rubinstein 1982, Powell 1987, Alesina and Drazen 1991).

The example of the series of wars between Israel and Egypt serves to illustrate how Brams (1994) applies the theory of moves in dynamic contexts. I first discuss Brams’s use of the case, then show how the theory of moves gets Brams into conceptual difficulties, and finally show how game theory can do a better job of explaining the events as Brams has reconstructed them.

Brams (1994) assumes that the central puzzle about the Arab-Israeli conflict was that it took so long for Israel to get to Camp David, that is, to make the concessions necessary to appease Egypt and end the cycle of fruitless conflict. He treats the Arab-Israeli conflict as a case of what he calls a “magnanimity game,” in which a victorious power can choose a far-sighted policy of making concessions to the vanquished or a short-sighted policy of preparing for the next war. Meanwhile, the vanquished can cooperate with the victor or resist. Brams places certain restrictions on the preference orderings in the magnanimity game and, in his discussion of the Arab-Israeli conflict, further limits himself to consideration of the following three situations, which he calls games 33, 34, and 35 (Figure 7).

Notice that in each of these games, Egypt has a dominant strategy of cooperating (it is, after all, very likely to lose if it chooses to fight) and that the unique Nash equilibrium in each case (underscored) is DC (no concessions, cooperate, which Brams labels “status quo”). Brams notes that in each case, Israel gets its best outcome and Egypt its second or third best, so he labels Israel “satisfied” and Egypt “dissatisfied.” He con-

7. Formally, any vector of payoffs that Pareto-dominates the worst vector of payoffs that the players can impose on each other as an equilibrium can be sustained as an equilibrium in a repeated game (Fudenberg and Maskin 1986; Fudenberg and Tirole 1991).
cludes, therefore, that game theory is not very useful in this case since it predicts neither war nor the ultimate resolution of the conflict by Israeli concessions. Nowhere is the Nash equilibrium of unilateral Egyptian cooperation the outcome that we observe. The static version to the theory of moves is not able to do any better, however. The outcome (no concessions, cooperate) is also the unique nonmyopic equilibrium if play begins at the status quo. Consequently, Brams turns to his alternative rule, which allows players to cycle through the game matrix indefinitely until one compels the other to stop by exerting moving power.

According to Brams’s (1994) rules, in each of the games in Figure 7, the player who is allowed to move next is always able to improve his or her position by cycling in a clockwise direction. First, Egypt shifts from cooperation to noncooperation, confident that this will induce Israel to shift to making concessions because this can only improve its payoff. Once Israel makes concessions, Egypt becomes more conciliatory, securing its first best outcome. After Egypt has begun cooperating, however, Israel no longer faces any incentive to make concessions, so it returns to the status quo. Egypt again shifts to noncooperation, and the cycle repeats itself. The only way to resolve such an impasse, for Brams, is to give one of the players moving power, which, in effect, is the ability to hold out longer than one’s opponent, who must eventually give up and stop moving. In this example, if Egypt had moving power, it could eventually force Israel to choose between the two outcomes where it has the next move, which are

<table>
<thead>
<tr>
<th>#33</th>
<th>Egypt</th>
<th>#34</th>
<th>Egypt</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C</td>
<td></td>
<td>C</td>
</tr>
<tr>
<td>Israel</td>
<td>D</td>
<td>1,2</td>
<td>Israel</td>
</tr>
<tr>
<td></td>
<td>3,4</td>
<td></td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>4,3</td>
<td>2,1</td>
<td>D</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>#35</th>
<th>Egypt</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C</td>
</tr>
<tr>
<td>Israel</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>3,4</td>
</tr>
<tr>
<td></td>
<td>1,3</td>
</tr>
<tr>
<td>D</td>
<td>4,2</td>
</tr>
<tr>
<td></td>
<td>2,1</td>
</tr>
</tbody>
</table>

Figure 7: The Arab-Israeli Conflict
the upper-right-hand and lower-left-hand corners of the matrix. Since Israel prefers the lower-left-hand corner, it eventually makes concessions from sheer exhaustion.

At first blush, this seems to capture the facts of the case. There were repeated clashes between Israeli and Egyptian forces in 1948, 1967, and 1973, with numerous bouts of saber rattling in between. Eventually, the superior Israeli forces withdrew from the Sinai to make peace, not because Israel feared defeat in another war but because Egypt appeared to be prepared to fight hopeless wars indefinitely. However, what is the substantive interpretation of the model’s prediction that Israel should make concessions whenever Egypt toughens its stance and then withdraw them whenever Egypt becomes cooperative? This sort of cycle never occurred, and it is hard to imagine that it could have. Instead, the Israel-Egypt face-off would be better described as a stalemate or war of attrition, punctuated by bursts of violence as the Egyptian side sought to exploit tactical surprise to win a total victory. Nor does Brams’s model capture the psychology of the confrontation, which was dominated by mistrust of each other’s motives and uncertainty about relative capabilities. In short, even if we accept one of the three preference orderings that Brams proposes, the dynamic story that he tells does not fit the story that he tells about the Arab-Israeli conflict.

Brams proposes two possible interpretations of cycling. First, it might be regarded as a hypothetical mental exercise, analogous to the trees of moves in a chess game that are left unplayed or to the off-the-equilibrium-path strategies and beliefs that play such an important role in a game-theoretic model. Egypt and Israel anticipate the possibility of cycling, Israel recognizes that Egypt has moving power (that it will eventually prevail if the cycle is played out), and consequently Israel never seizes the opportunity to move back to the status quo. If this is the case, however, how are we to explain 31 years of conflict and repeated outbreaks of war? If the cycling is purely hypothetical, then it does not help us to make empirical predictions or resolve the puzzle that Brams originally set out: why did it take so long to get to Camp David? On the other hand, the cycles may be regarded as actual moves and countermoves by the players. However, this confronts the issue raised above: the predicted moves are empirically incorrect and even extraordinarily unrealistic. Even if the moves are regarded as hypothetical, moreover, the fact that they are so unrealistic forces us to stretch our credulity to believe that decision makers acted as they did because they anticipated precisely those responses.

At this point, we are in a position to evaluate Brams’s third claim about the theory of moves.

Claim 3: The theory of moves makes theoretical progress in understanding repeated games by exploring dimensions of power—threat power and moving power—that game theory cannot grasp. I argued above that threat power and moving power are both loose ways to discuss phenomena that formal game-theoretic analysis makes more precise. Threat power is an analogy to the finding that cooperative strategies can be supported by threats in infinitely repeated games; moving power is an analogy to repeated-game models of bargaining. In discussing Brams’s (1994) use of moving power to explain the course of the conflict between Israel and Egypt, I argued that the example seems contrived because the strategies that the theory of moves ascribes to the players are so unrealistic. The deeper theoretical problem, however, is that Brams
treats the central issues in his bargaining model informally. First, Brams leaves unspecified exactly why one player is compelled to stop cycling. Presumably, cycling is costly in some sense, either because of the delay in reaching resolution of the issues or because of increased risks of conflict. However, Brams never spells out the costs and never incorporates them into his model. One player is presumed to be privileged, with a greater ability to bear these costs, and this player has moving power. The determination of which player has this advantage, however, is essentially arbitrary. Instead of investigating the dynamics of resolve in bargaining, as game theorists have done by explicitly modeling the cost of delay, he simply assigns greater resolve to one of the players, grants it moving power, and treats that as an explanation. In practice, the analyst examines a historical case, finds a way to fit it into the pattern of some $2 \times 2$ matrix, and assigns moving power to whoever must have had it to get the desired outcome.

Second, Brams (1994) has to appeal to incomplete information to explain lengthy delays in resolving conflicts; indeed, if the actors are rational, any delay at all is inexplicable in a model with full information. In the Egypt-Israel confrontation, he argues that Israel was uncertain about Egyptian resolve and that Egypt consequently had to bear substantial costs over a long period of time to prove to Israel that it was committed enough to its objectives to continue the struggle indefinitely. This, indeed, is consistent with accounts of Sadat’s motivations in starting the Yom Kippur War in 1973. In the context of Brams’s game, however, there is nothing for Israel to be uncertain about. The preference orderings and capabilities of the players are common knowledge. Similarly, Brams argues that Egypt sent a signal of resolve by moving toward confrontation; but because this is a costless action in the game that he uses as a model, it is not an action that could be used as an informative signal. When Brams is finished, all that is left of signaling games and games of incomplete information are vague analogies. This is unfortunate because these are concepts that game theorists have made very precise and powerful and with which they have made some of their greatest advances in the past two decades.

8. If we treated the players’ payoffs as cardinal utilities, we might suppose that there was uncertainty about their discount factors, but since Brams only allows payoffs to be determined at the end of the game and without discounting, this does not help.

9. Brams (1999) revisits this issue in his most recent article on the theory of moves in which he analyzes two earlier crises between Egypt and Israel: the 1960 Rotem crisis and the 1967 crisis that led to the Six-Day War. He assumes a different set of 12 preference configurations than those in his 1994 book and also differs from Mor (1993). In addition, he makes several amendments to the rules of the theory of moves. He attempts to capture the dynamics of crisis escalation as cycling through four states: status quo (no one mobilizes), unilateral escalation by Egypt, crisis (both mobilize), unilateral demobilization by Egypt, and back to status quo. When he describes the events of May and June 1967, however, they do not fit this structure. Brans describes these events in terms of three stages: the Egyptian mobilization, the Egyptian request to withdraw UN ceasefire monitors, and the Egyptian decision to close the Strait of Tiran. According to Brans, Israel did not escalate after the initial Egyptian mobilization (although it did mobilize some forces) so play ends at unilateral escalation. However, when Brans analyzes the next stage, he assumes that play again begins from the status quo, which was unsatisfactory for Egypt because “its expectations have now been raised by its partial success” (n. 636). In the second stage, Brans emphasizes that Nasser hoped to avoid war but was emboldened by Israel’s mild response to his first challenge and demanded that the UN remove its troops. Israel mobilized for war but exercised diplomatic restraint, and again Brans returns the disputants to the status quo. In the third stage, Nasser blockaded the strait and signed a mutual defense pact with Jordan, which provoked Israel to launch a preemptive war that returned the situation not to the status quo ante but to a situation much less advantageous for Egypt and its allies.
APPLICATIONS OF THE THEORY OF MOVES

Several studies have recently appeared in peer-reviewed journals that use Brams’s (1994) method to interpret cases in international relations. This section will show that (1) the predictions that the authors deduce from the theory of moves can be replicated in game-theoretic terms by an extensive-form game; and (2) the models that the authors use do not, in fact, fit the stories that they tell about particular cases, and game theory can provide models that fit those accounts better. The cases concern economic sanctions, the repatriation of refugees, and the initiation of international crises.

ECONOMIC SANCTIONS

In an article that appeared in *International Interactions* in 1996, Marc Simon asserts that the theory of moves “provides a better conceptual account of the dynamics of sanctions disputes over time than traditional game theory” (p. 203). Simon applies the theory of moves to economic sanctions by constructing a general model that treats sanctions episodes as $2 \times 2$ games and makes assumptions that narrow the possible range of preferences to 11 configurations. In particular, he assumes that targets of sanctions have a dominant strategy not to comply: they prefer not to comply if the sender does not impose sanctions, and they prefer not to comply if the sender imposes sanctions regardless of what they do. He notes, correctly, that if these models are treated as normal-form games, game theory predicts that sanctions never succeed. Why, then, do nations often impose sanctions? He concludes that the theory of moves provides a more realistic model of sanctions since several of his preference configurations have nonmyopic equilibria in which sanctions lead to compliance.

Simon’s (1996) game-theoretic opponent is a straw man. The problem is that single-shot, simultaneous games are not good models of sanctions. Using a simultaneous game as a model means that countries have to choose whether to comply without knowing whether sanctions have been imposed and whether to impose sanctions without knowing whether the target has complied. Furthermore, because this is a single-shot game, compliance and sanctions are both irreversible outcomes. The model rules out the following strategy: impose sanctions and lift them only if the target complies. This, however, is the strategy that senders almost always choose. Under the circumstances, we should not be too impressed by the fact that this model fails to predict success.

Simon (1996) illustrates the use of the theory of moves using two cases of U.S. sanctions: against Vietnam (1975-94) and Haiti (1985-94). He represents the U.S.-Vietnam case after the end of the cold war using game 27 (Brams 1994) and posits that

Contrary to Brams’s theoretical claim, there is no cycling in his story; Egypt does not advance and retreat. Instead, there are three stages of a repeated game in which Egypt advances fait accomplis, and Israel decides whether to attack. In fact, the model that best fits the story is Powell’s (1987) model of crisis deterrence under incomplete information. Brams states that Nasser thought his actions increased the chance of war from 20% after the withdrawal of UN troops to 50% when he closed the strait and to 80% after signing the Egypt-Jordan defense pact, which is consistent with a model in which states deliberately take actions that increase the risk of war. In addition, he emphasizes that Nasser miscalculated Israeli resolve, which points to a central role for incomplete information in his explanation. However, neither chance moves nor incomplete information has any place in his model.
play begins in the lower-right-hand cell (Figure 8). According to the theory of moves, if play starts in the lower-right-hand corner, Vietnam will change its strategy to cooperation, and the United States will drop the sanctions. This is what actually happened.

This argument is equivalent to proposing an extensive-form game as a model of the U.S.-Vietnam sanctions episode (Figure 9). Left-hand options signify noncooperation for Vietnam or the imposition of sanctions for the United States, and right-hand options signify cooperation or the lifting of sanctions. The sequence of choices is identical to the argument that Simon (1996) presents. Vietnam can choose to resist sanctions or comply; the United States can lift sanctions or continue them; Vietnam can decide whether to continue to comply; if it fails to comply, the United States can reimpose sanctions. The unique subgame perfect equilibrium is the bold sequence leading to (4,3).

On the other hand, is this a good model of the case? Simon (1996) emphasizes that there was a lengthy period during which the Vietnamese attempted to convince the United States that they were complying in good faith on the POW/MIA issue, and the United States continued to impose sanctions and block Vietnam’s access to international financial institutions. The empirical puzzle is why the sanctions were maintained for so long in spite of the fact that a Pareto-superior alternative existed. This is something that the theory of moves (or the extensive-form version of the same model) cannot explain because these models do not expect the shifting of strategies to occur in

10. Brams (1994) and Simon (1996) seem to disagree about the rules of the theory of moves at this point. The interpretation just given is consistent with Brams’s rules, to which Simon claims to adhere, and with Simon’s story, in which we arrive at game 27 after the United States has been imposing sanctions for a number of years. However, Simon claims that he starts play in the upper-right-hand corner. From that point, according to the theory of moves, play ought to cycle (if cycling is permitted) or go to (3,2) (if cycling is prohibited). Brams could get the prediction Simon makes either by starting play at the lower-right-hand corner or allowing cycling and giving the United States moving power. Simon gets his solution by prohibiting the game from cycling back to its starting point, which is a restriction that Brams does not make.
real time. No payoffs are supposed to accumulate as the players move around the game matrix, yet it is clear that the Vietnamese did bear costs during the years in which they were subjected to trade sanctions. Furthermore, there is nothing for the United States to doubt in the model: all of Vietnam’s actions and preferences are transparent. A better model would be a repeated game with incomplete information in which the choices of the players about compliance and sanctions impose real costs, and the United States is
uncertain whether Vietnam intends to comply. In such a model, the United States
might subject Vietnam to a lengthy testing period before dropping sanctions.\textsuperscript{11}

Simon (1996) uses game 48 to analyze the U.S.-Haiti case (Figure 10). He argues
that the difference between the two games is that the United States preferred unilateral
cooperation to unilateral defection in the Haiti case and had the opposite preference
ordering in the Vietnam case because the U.S. lobby in favor of sanctions was much
stronger. When treated as a normal-form game, game 48 has the same Nash equilib-
rium as game 27: repress, impose sanctions.

According to the theory of moves, game 48 is quite different from game 27: both of
the cells in the top row are nonmyopic equilibria from some starting points. If the theo-
rist allows the game to cycle, it ends at one of these two points, depending on who has
moving power. Simon (1996) argues that this is what happened in the U.S.-Haiti case.
From the fall of Duvalier in 1986, he traces four cycles in which the United States
imposes sanctions, the military permits civilian government, the United States drops
sanctions and extends aid, the military returns to repression, and the United States
reimposes sanctions. Finally, in 1994, the United States prevails, and the military
retires from politics.

How does Simon (1996) explain this outcome? In the theory of moves, cycling is
resolved by invoking moving power, but the assignment of this power to one player or
the other is based on a subjective judgment. “Given its immense military power and its
history of strong ties and influence within the Haitian military,” Simon writes, “it is
reasonable to assume [italics added] that the U.S. believed that it had the capability to
influence the Haitian military to stop cycling at some point.” (p. 223) This is troubling.
Do we want to assume that the state with the most military capabilities always prevails
in tests of will? This argument, furthermore, conflicts with Brams’s (1994) rationale
for assigning moving power to Egypt in the Arab-Israeli war of 1973 because, although
weak, Egypt enjoyed popular support for its cause. If we grant the assumption that the
United States had moving power in this case, however, why did the cycles continue for
8 destructive years? Simon suggests that there may have been uncertainty about who
possessed moving power: “In fact, the ambiguity of U.S. support for sanctions, weak
support for Aristide, and continued clandestine support for the Haitian military may
have led the Haitian military to mistakenly conclude that it had moving power” (Simon
1996, 223). He goes on to speculate that the Haitian military might revive its ambitions
after the U.S. occupation force withdrew, reopening the question of who really had
moving power. In fact, he can only infer the possession of moving power from the out-
come, which makes the argument circular.

Is there a better model for this case? One approach would be to model this situation
as a two-level game between the United States, the current Haitian military strongman,
and the Haitian military. This would help to capture the intuition that it makes sense for
individual officers to contemplate a coup for the sake of personal advancement, but the
military may withdraw its support from them if the United States imposes severe costs.

\textsuperscript{11} For example, if Vietnam is one of two “types” (cooperative, uncooperative), the testing period may
allow the cooperative types to separate themselves from the uncooperative types, who are less willing to bear
the costs of unilateral cooperation.
When outcomes appear irrational, it generally makes sense to look for more actors. Without giving up the unitary actor assumption, however, it would be possible to explain this sequence of events in terms of a repeated game with incomplete information about U.S. intentions. In this model, the United States might have to impose sanctions repeatedly to signal its resolve. In the end, of course, the United States had to signal willingness to do more than impose sanctions to get the Haitian military to back down.

**REPATRIATION OF REFUGEES**

In an article that appeared in the *Journal of Conflict Resolution* in 1996, Lester Zeager and Johnathan Bascom apply the theory of moves to the problem of international refugees. Zeager and Bascom seek to explain the difficulties the United Nations High Commissioner for Refugees (UNHCR) has had in repatriating refugees, and they develop four different 2 × 2 games based on different assumptions about the players’ preferences. They argue that “the theory of moves is better suited [than game theory] for understanding efforts to achieve repatriation agreements in actual refugee crises” (p. 460) and that “concepts such as moving power and threat power in TOM can uncover some of the reasons for success and failure in actual repatriation negotiations” (p. 483).

To illustrate their method, consider their treatment of the Eritrean case. After the Ethiopian civil war, the UNHCR sought to repatriate some 500,000 Eritreans who had fled to Sudan, and Eritrea insisted that it was unable to absorb such numbers of refugees unless the UNHCR quadrupled the amount of resettlement aid that it proposed to offer. Zeager and Bascom (1996) model this as a case of an “ambivalent government and apathetic donors,” which is game 47. In this game, the UNHCR actually prefers the outcome of continuing to assist the Eritrean refugees in Sudan to the outcome of repatriating them to Eritrea at higher cost. Zeager and Bascom justify this assumption.

<table>
<thead>
<tr>
<th></th>
<th>Haitian</th>
<th>Military</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States</td>
<td></td>
<td></td>
</tr>
<tr>
<td>~ sanctions</td>
<td>4,3</td>
<td>2,4</td>
</tr>
<tr>
<td>sanctions</td>
<td>1,1</td>
<td>3,2</td>
</tr>
</tbody>
</table>

*Figure 10: U.S.-Haiti Sanctions Game*
by arguing that the UNHCR actually represents a coalition of donor countries and that
the one-time costs of repatriation are much higher than the average annual costs of
maintaining refugees in exile (Figure 11). If this is treated as a normal-form game,
there is no Nash equilibrium in pure strategies (there is, of course, a mixed-strategy
equilibrium). The authors claim that since there is no pure-strategy equilibrium, game
theory does a poor job of explaining the outcome of this case. Again, however, a
one-shot, simultaneous-move game is prima facie a poor model of what is essentially a
bargaining situation, so this appears to be an irrelevant criticism of game theory.

In the theory of moves, this is a game that cycles in a clockwise direction. Starting
from the lower-right-hand cell, UNHCR extends aid to Eritrean refugees in Sudan,
Eritrea agrees to repatriate them, the UNHCR withdraws its assistance, Eritrea with-
draws the offer to repatriate, the UNHCR again offers aid, and the cycle continues.
Either cell in the left-hand column can be a stable outcome depending on who enjoys
moving power. Zeager and Bascom (1996) argue that the UNHCR had moving power
in this episode, so their model would predict the lower-right-hand outcome (assistance
but no repatriation).

The authors argue that this fits the Eritrean case well. When the UNHCR initially
offered aid, Eritrea responded by allowing the first group of refugees (about 70,000) to
be repatriated but then demanded more money before the program could be expanded.
The international donors refused to support the UN appeal to provide additional funds,
however, and Zeager and Bascom (1996) interpret this as a shift in strategy. In conse-
quence, only another 22,000 refugees were resettled, and the outcome settled down
into the lower-left-hand cell, where the UNHCR continues to provide aid to expatriate
refugees. “This scenario fits remarkably well with the prediction of our model when
UNHCR or, more generally, the party aligned with donors, has moving power”
(Zeager and Bascom 1996, 479). However, assigning moving power to the UNHCR
seems arbitrary: it is another case of selecting the model based on the outcome, which
leads to circular reasoning. In fact, it is unclear from the story the authors tell why we
need a model that involves cycling and one party outlasting the other. Neither side
compromised its bargaining principles: the Eritreans gave the donor countries a
take-it-or-leave-it offer, and the donors rejected it. A better model would emphasize
the role of incomplete information in a sequential bargaining game. Figure 12 is one
such game.

Branches leading to the right indicate cooperative moves by each party, and
branches to the left indicate uncooperative ones. Eritrea’s choices are a high or low
demand for international aid in return for resettlement; the UNHCR chooses to accept
or reject Eritrea’s proposal. In this model, Eritrea is uncertain whether it faces a forth-
coming international donor community (right-hand branch) or a reluctant one. If the
international community is forthcoming, Eritrea gets its best outcome (4,3) by striking
a tough bargaining stance because either offer will be accepted; if the international
community is reluctant, the high demand will be rejected, leading to (1,3), and the low
demand will be accepted (3,4). Eritrea’s optimal strategy depends on its assessment of
what type of donor community it faces. If there is enough uncertainty, however, that
optimal strategy may sometimes lead Eritrea to demand too much and the donor com-
munity to reject its proposal. Note that in this model, we do not have to make the offers and contributions exogenous, as Zeager and Bascom (1996) did, and we do not have to make the strong assumption that the international community prefers to see the Eritrean refugees remain in Sudan. By modeling the problem of bargaining under incomplete information explicitly, we arrive at a more realistic model. Of course, I hasten to add that the model that I propose is no less circular than Zeager and Bascom’s. After all, this is a model-fitting exercise, and to determine whether such a model is useful, we would have to test it against data that were not used to generate it. The point here is simply that the more flexible tools of game theory can provide a model that fits the story that Zeager and Bascom tell about their case better than the model that they derive from the theory of moves.
CRISIS INITIATION

Perhaps the most interesting article in this literature is a piece by Ben Mor that appeared in 1995 in the *Journal of Theoretical Politics*. Mor uses the theory of moves to develop a fairly general model of international crises, deduces from the model conditions for crisis initiation, and investigates the effects of misperceptions of opponents’ preferences. Mor adopts the convention that games stop after a single cycle, but he makes the significant change to the theory that the valuation of the starting point may be different after the cycle is complete. His model of international crises is illustrated in Figure 13.

C is defined as a cooperative move, such as compromise or backing down from a military conflict, and D is defined as an aggressive move, such as making a demand or escalating the conflict. Play begins in the upper-left-hand corner, the status quo (SQ); however, if play cycles back to this point, it has a different value, which is associated with compromise (CC). Mor makes the sensible assumption that both players prefer the SQ, CC, and victory to acquiescence; otherwise, preferences can take any order. By assumption, the initiator is the row player, so movement around the matrix is counterclockwise.

Mor (1995) analyzes the possible conditions for crisis initiation by backward induction, and in this case the analysis is more complex under the theory of moves than under game theory because Mor has to trace out the moves and countermoves generated by every possible combination of preferences. The equivalent extensive-form game is much more intuitive (Figure 14).
Left-hand branches signify aggressive moves; right-hand branches signify conciliatory moves. Figure 14 makes it apparent, for example, why Mor (1995) concludes that the initiator may initiate a crisis if it prefers CC to SQ but only if the defender prefers CC to CD; most of the results are of this sort. Mor’s main results regard satisfaction with the status quo and misperceptions. First, he concludes that satisfaction with the status quo plays a key role in determining whether crises are initiated, and, again,
Figure 14 shows why this is the case. The decision to initiate is always a weighing of the expected value of a crisis against the value of the status quo, so if the initiator places a low enough value on SQ, it will always choose a crisis over SQ. Second, Mor notes that misperception of the opponent’s preferences can lead either to unwarranted initiation (and sometimes war) and to unwarranted failure to initiate crises. Third, he finds that misperception has the least influence when the status quo has a low value; conversely, misperception is most likely to be harmful when the status quo is most valuable. Again, Figure 14 shows why this should be the case: all decisions to initiate are weighed first against the status quo. If the status quo is bad enough, the initiator always initiates, regardless of the type of the defender; if it is good enough, the initiator only initiates when the prospects are favorable, so its decision is more sensitive to perceptions of the opponent’s type. Note that none of these conclusions are dependent on the theory of moves; they all follow from the extensive-form game that represents the series of choices that Mor derived from the theory of moves.

The extensive form of the crisis initiation game seems to be a reasonable way of representing crises; however, it is arbitrary. An extensive-form game embodies a theory about what the strategic situation looks like. Leaving the structure of the extensive form up to an arbitrary process such as the theory of moves means generating a theory randomly. Mor (1995) would have generated a different extensive form, for example, had he started from a different cell in his matrix or had he followed the ordinary theory of moves convention that a cell cannot change valuation in the course of a game (SQ/CC). To illustrate the significance of these choices, it is useful to compare the crisis initiation game to a prominent alternative model, the international interaction game proposed by Bueno de Mesquita and Lalman (1992). The simpler model proposed by Mor eliminates a number of branches of the tree with quite dramatic consequences. First, in Mor’s theory, the status quo is an outcome that can always be imposed by the initiating state; in the Bueno de Mesquita and Lalman model, both states have to choose the status quo for it to persist. The consequences of this simple modeling choice drive the divergent results of the two models. Bueno de Mesquita and Lalman discover that the valuation of the status quo does not influence choices about war because in their model, the status quo is only accessible when a negotiated outcome is possible in any case, and a negotiated outcome is preferable to war. Second, in Mor’s model, the choice of war precedes the choice of compromise so it is impossible to retain the option of fighting and force negotiations from a position of strength. Likewise, it is impossible for the defender to offer negotiations to prevent war unless the defender prefers negotiations to victory. In Bueno de Mesquita and Lalman’s international interaction game, by contrast, the chance to fight a war remains open as long as the crisis continues, so a country with a credible threat to fight can force negotiations rather than capitulate.

Mor’s (1995) work on crisis initiation represents the very best that can be hoped for from the theory of moves: it generates an arbitrary extensive form that actually corresponds reasonably well to the phenomenon of which it is intended to be a model. Whether the international interaction game or the crisis initiation game turns out to be a better model of international crises is a matter for empirical testing to determine. At this point, however, it is important to note that the function of the theory of moves is
simply to generate an arbitrary extensive form. The fact that this model was generated by (a modified version of) Brams’s (1994) rules does not make it intuitively any more appealing as a model of conflict than any other randomly drawn extensive form. Nor are the modeling choices embodied in those alternative extensive forms insignificant, as the comparison with the Bueno de Mesquita and Lalman (1992) model demonstrates.

CONCLUSIONS

The extensive form matters. The most fundamental reason why the theory of moves is an intellectual dead end is that preference orderings alone do not convey enough information about a game to derive general conclusions. There have been a few general conclusions based on preference orderings: for example, the minimax theorem solves all constant-sum games. In general, however, to analyze a game, it is necessary to know the order in which moves occur and what information the players have. All that the games operated on by the theory of moves share with their analogous normal-form games is the order of the actors’ preferences, so there is no reason to suppose that any of the results of the theory of moves will apply to situations that are characterized by those normal-form games. The theory of moves can say nothing about prisoner’s dilemma, for example, because prisoner’s dilemma is a simultaneous-move game.

The theory of moves does not generate any insights that are not captured by game theory. Because the series of choices generated by the theory of moves can be translated into extensive-form games, all the theory of moves contributes is a redundant vocabulary. The theory of moves does not treat the extensive form as endogenous, as its adherents claim; rather, each game has an exogenously fixed extensive form that is determined by the analyst’s choice of a starting point and a first mover. Reanalyzing articles that use the theory of moves shows that when it appears to account for empirical cases, this is because the extensive-form games that it generates happen to fit those cases plausibly. More frequently, the cases have to be stretched to fit a model that is not quite right. In these cases, game-theoretic models can always be customized to capture details of strategic interaction, timing, and information more precisely.

The theory of moves represents backsliding, not progress, in the effort to understand strategic interaction. Normal-form, extensive-form, and repeated games and games of complete and incomplete information provide a powerful language for translating strategic situations into formal notation. Using these tools, game theorists can make discoveries by developing rigorous mathematical proofs. For example, game theory has proven that repetition can mitigate the problems raised by opportunistic behavior; the theory of moves can only assert this as an assumption because it has no way of analyzing repeated games. The effort to prove this has led game theorists to make important discoveries about strategic interaction. We have learned, for example, that there have to be strategies to punish the guardians in case they should happen to shirk their responsibilities; we have learned that agents must be patient enough to prefer long-term gains over short-term temptations; we have learned that some cooperative equilibria
exhibit evolutionary stability, and others do not. The theory of moves is unable to form-
italize its arguments and consequently holds no similar promise of generating theoretical progress.

REFERENCES


