Mershon International Studies Review (1995) **39**, 19–52

Modelling Decisions in International Relations: Game Theory and Beyond

PETER G. BENNETT

Department of Management Science, Strathclyde Business School
University of Strathclyde

Renewed interest in Game Theory and its extensions comes at a time when richer models than ever before are on offer. This paper starts with a non-technical introduction to game models, arguing that they can best be used to provide an understanding of decisions, rather than prescriptions. It then examines some common criticisms of the approach, in particular those surrounding the concept of "rational choice." Extensions to the basic model are then illustrated: for example, approaches stressing differences in perception or introducing "internal" and "external" games. Rather than being restricted to rational choice in a narrow sense, these approaches offer a richer picture of decision-making as "reflective choice." More radical possible developments include replacing the "game" metaphor by that of "drama," and combining formal analysis with empirical theory to create "knowledge-based" modelling methods.

Interest in the use of Game Theory to model international politics has alternated between periods of high hopes and great expectations and periods of disfavor. The one constant has been controversy. When John Von Neumann and Oscar Morgenstern's classic *Theory of Games and Economic Behavior* appeared in 1944, some were inclined to see it as the ultimate theory of all conflictual and competitive decision-making—whether in economics, politics, or international relations (IR). Others have always been ready to attack Game Theory as unrealistic, unhelpful, or downright pernicious. Controversy was heightened by the theory's association with the politics of the early Cold War. Though some of the issues raised then remain highly relevant, it is now easier to disentangle the essentials of the theory from specific policy debates now long since defunct. Game Theory has come somewhat back into vogue in the last decade. While several lines of research hold out the promise of more realistic models than those offered in the 1950s and 1960s, controversy remains, and many are ready to dismiss anything containing the word "game."

It is therefore appropriate to look at game-based modelling in general and to assess both its promise and its limitations. This review does so through five main parts.

Published by Blackwell Publishers, 238 Main Street, Cambridge, MA 02142, USA, and 108 Cowley Road, Oxford OX4 1JF, UK.

1. The first section provides a non-technical introduction to the idea of a "game," gives two simple illustrations, and discusses the interpretation and use of game models in general.
2. The second section examines some common objections to game-based thinking in IR—both those charging that game models over-simplify reality, and those alleging that Game Theory has fundamental philosophical flaws. Particular attention is given to the notion of "rationality" assumed within the theory.
3. The third part outlines extensions to the basic "game" idea so as to suggest the repertoire of models now available. For example, some approaches stress that conflictual parties may define the conflict quite differently; other models seek to deal with the dynamics of conflict. This section concludes with some further comments about modelling methodology, especially on the need to choose which forms of complexity to include.
4. The next section discusses three lines of further development:
   - a more radical re-casting of the underlying theory, replacing the "game" model with that of "drama" and combining analysis of both rationality and emotion;
   - combining game (or drama) models with some of the "knowledge-based" methods associated with Expert Systems approaches;
   - using the development of flexible, user-friendly computer software to open up new possibilities for both decision support and field research, with analysis done in collaboration with actors directly involved in the issues.
5. The final section illustrates some of these developments by reference to some research on the conflict in Northern Ireland.

Though partly *about* Game Theory, this review is not intended as an essay *in* Game Theory. It discusses its assumptions and uses, rather than detailing, the deductive theory itself. Readers interested in the latter may find it helpful to peruse such texts as Shubik (1984), Thomas (1986), or Ordeshook (1986). Those already familiar with Game Theory may still find the developments and suggestions outlined here of some interest. The overall message is that, though some criticisms of Game Theory are based on misunderstanding, and others can be met through extensions of the basic model, the approach does have its limitations. There is a strong case for integrating Game Theory with other perspectives, as part of a more eclectic methodology. Above all, care is needed to avoid extravagant claims. Models can enhance understanding, but one should not expect them to reveal the right answer to complex issues.

The "Game" Model

Since many people are repelled by the language of Game Theory, it is important to note straight away that the term "game" is not meant to imply any sense of trivialization. A game in this sense is simply a situation in which there are two or more parties who can affect what happens, all of whom are pursuing their separate aims. Such situations provide particular challenges, both for decision makers and for students of choice. Because no actor has complete control over events, each needs to take account of the others' possible actions. Suppose actor A's most advantageous course of action depends on what actor B decides to do, and vice versa. If both realize that this is so, A will try both to anticipate and to influence B's choices, knowing that B is trying to do the same in reverse. Thus even if A and B never meet, their decisions interact, and they will find themselves in an outcome

dependent on both their choices. Such situations typically bring forth possibilities for mutual threats, deceit, bluff, and counter-bluff. But these are still only one side of the story. Conflict is seldom absolute. Some common interests can almost always be discerned, so joint gains can be made by cooperating—or at least keeping conflict within limits. Thus interactive decisions are also concerned with cooperation and "collaborative advantage" (Huxham and Macdonald, 1992) and with promises as well as threats. Where aims partially diverge—as they usually do—conflict and cooperation become two sides of the same coin and cannot properly be dealt with separately (Lax and Sebenius, 1986).

Game models are intended to provide an analytical guide through this maze. There are various ways of representing a game, perhaps the most intuitive being as a tree of moves, in which each branch, or move, is under the control of a particular player, and the moves available at any point may depend on those already made. One can visualise the game of chess, for example, as a fantastically-complicated tree of possible moves. This *extensive form* model will be discussed later, but a simple tree of moves is illustrated in Figure 1. This two-player case has just seven possible outcomes (end points of the tree).

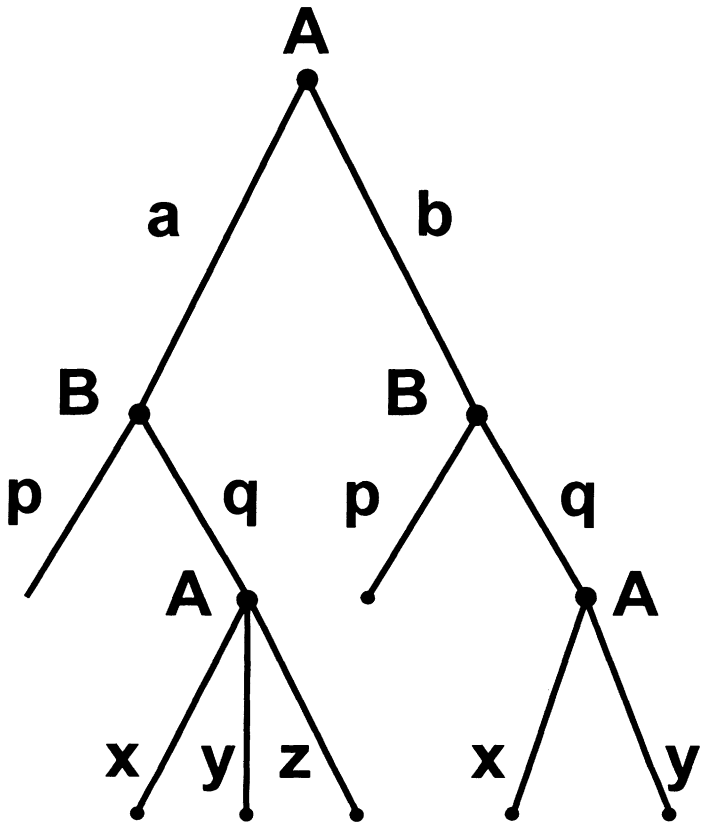

FIG. 1 Structure of a simple game tree (extensive form model)

Player A moves first, choosing between moves a and b. B then chooses p or q. If B chooses q, A has a further choice either between x, y and z (if initial choice was a), or just between x and y (if initial choice was b).

Much of Game Theory—at least until recently—has been based on Von Neumann's notion of "strategy." Technically, a strategy is a complete recipe of actions, specifying what is to be done in every possible circumstance that might be encountered. In the case shown in Figure 1, for example, *one single* strategy for player A is to choose option "a" initially, then "y" if B chooses "p," but "x" if he chooses "q." In this way, it was argued, a series of choices could be telescoped into one single choice of strategy without any loss of generality. It was this simplification that paved the way for the first general analyses of games. In this strategic, or *normal form,* a game consists of the following:

- Two or more *players,* that is, the interested parties. For the purposes of the model, players can be individuals, groupings such as committees or cabinets, or entities such as corporations or nations.
- A set of *strategies* for each player, representing possible courses of action, each including all necessary conditional clauses.
- The set of possible *outcomes* of the game, given by considering all the combinations of strategies that the players could choose.
- A set of *preference functions* specifying how good or bad each outcome is for each player. Game Theory is often thought to require quantification of preferences by means of *utility scales.* However, many models assume only that one can define a *preference order* for each player, that is, outcomes are simply ranked from best to worst.

There are many types of games within the ambit of the theory. Games can have two, three, or many players. They can be played with or without communication. They may be played just once, or repeated many times. A game may include chance events as well as deliberate moves; this is modeled by introducing "nature" as a separate player, presumed to have no preferences but to act in accordance with certain probabilities. Co-operative Game Theory allows players to form binding coalitions, although the more fully-developed non-cooperative branch of the theory does not. But the most fundamental distinction among types of games concerns the relationship between the players' aims. At one extreme, their preferences may be exactly opposed, so that any gain by one must be at the other's expense. If preferences are expressed by utility scales, the total utility is always the same (in effect, the game is to divide a "cake of fixed size"). Hence these are known as constant-sum games. By convention, one usually sets the constant to be zero (measuring the losses and gains relative to an equal division), thus a *zero-sum game.* (With two players, the term is often used colloquially to cover any case in which one side's preference ranking is the opposite of the other's.) At the other extreme, the players' aims may exactly coincide. Their problem is then one of *pure coordination* in trying to achieve what they both want. Such games need not be trivial, and interesting work has been done on the origins of social conventions as solutions to pure coordination games. Between these extremes are games in which the players' interests are in partial conflict, so that they may be torn between conflict and cooperation. These are known as *mixed-motive games.*

Thinking carefully about a situation in terms of relevant players, strategies, outcomes, and preferences can be a useful exercise in itself, drawing attention to features that might otherwise be overlooked. For most Game Theorists, however, the main point is to bring some deductive analysis to bear. The first and simplest form of analysis assumes that every player knows what the game is (that is, has complete information about everyone's available strategies and preferences) but does not know what strategy has been chosen by the other player(s). The theory then provides conclusions about the decision problem. In particular, analysis de-

termines the existence of any *equilibria:* outcomes in which each strategy choice is a rational response to the other(s). Such choices are mutually reinforcing: anticipating the other players' decisions would not lead one to change one's own. Formally, there are various kinds of equilibria, the simplest being a *Nash Equilibrium,* in which no individual player can achieve a better outcome by changing strategy as long as all other players' choices are held constant. Other criteria allow for moves by coalitions rather than just individual players, or suppose that any change may set off a sequence of responses and counter-responses rather than holding other players' choices constant (Brams, 1985).

Game Theory began (and for many, continues) as a normative theory, the aim being to show how rational players *should* act. For some classes of game, this can be done quite convincingly. In particular, Von Neumann and Morgenstern offered a general analysis of two-player zero-sum games with fully-quantified utilities. Optimal choices were shown always to exist, provided players could choose probability "mixes" of strategy. Admittedly, a game might be too complicated to analyze completely—as with chess. But the theory could still throw an interesting light on what sort of solutions existed, and some non-trivial cases certainly can be solved in practice. Even then, however, the recommended strategies are only rational ways of playing against an equally rational opponent. They do not show how to exploit an opponent's mistakes. Much more importantly, zero-sum situations are fortunately rare in real life. (General Douglas Haig reputedly rebuffed criticisms of trench warfare by arguing that Allied troops outnumbered the enemy by half a million. So if they could kill each other at the same rate, he would eventually "win" by 500,000. Even in military circles, this is generally regarded as unimaginative; and if both sides prefer to avoid a bloodbath, the game is not zero-sum.)

Mixed-motive games are much closer to the real world of partial conflict, threats and promises, bargaining and negotiation, dispute settlement, and—arguably—conflict resolution. In these games, convincing normative solutions are more difficult to find. This has led to different ways of interpreting the theory, affecting the *types* of conclusions drawn.

Two Illustrations

Two well-used examples will serve to illustrate these ideas. The first is the game of Chicken. This has frequently been used to model situations of international brinkmanship such as the Cuban Missile Crisis, and such a model is represented by the normal form *matrix* in Figure 2. Each player's possible strategies are shown along one side of the matrix: one player controls the rows, the other the columns. (A three-player game would have a 3-D matrix, and so on.) Each cell of the matrix represents a possible outcome, and in each cell are the players' preferences. In this case, there are two players—labelled "White House" and "Kremlin"—each with just two choices—"back down" or "stand firm." The figures shown in the cells represent preference rankings from 4 (best) down to 1 (worst). (Those for "White House" are on the left in each pair.) Though the two sides clearly do not have the same preferences, this is not a straightforward zero-sum conflict. If neither side backs down, *both* end up with their worst possible outcome. Note that the model requires only preference orders. It does not depend on being able to quantify the desirability of the outcomes.

The model, of course, draws on an analogy with games of Chicken played in cars or on motorcycles by Hell's Angels and the like—the choice being whether to be a live but despised "chicken" by swerving first, or to risk disaster by driving on. However, the conclusions drawn can be applied to any situation—regardless of context—in which there are two sides with these preferences. There is a tricky

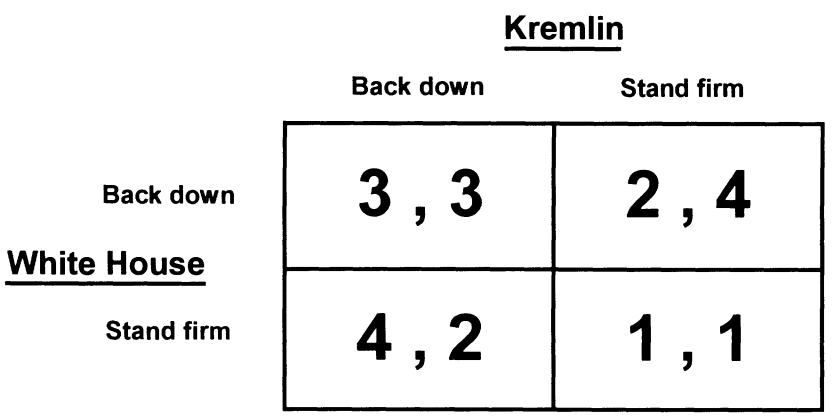

FIG. 2 Cuban missile crisis modelled as a game of chicken, showing matrix normal form model.

problem of choice here. If you assume that the other side will choose in the same way as yourself, then *either* choice must turn out "wrong." If she swerves, you should have driven on, and vice versa. Analyzing the game reveals two (Nash) equilibria: "4,2" and "2,4." That is, a "win" for either side can be stable, because the loser could only move to an even worse outcome. However, the game is symmetric between the two players, so there is no way of telling which side will establish a winning position. In practice, the model tells us to expect a race to establish commitment, each side trying to convince the other that it cannot or will not back down. Thomas Schelling (1960) provides a classic exposition of the tactics to be expected, many of which can indeed be observed in everyday life as well as in high politics.

The second game is known as Prisoners' Dilemma. The name derives from a story about two suspected criminals held in separate cells, each given the choice of squealing or maintaining silence. In IR, it is frequently used as a simple model of arms races (Rapoport, 1960), and in this form it is shown in Figure 3. The game involves two (in this case, fictional) nations, Andersland and Jedesland, each again having just two possible strategies, "arm" and "disarm." Four outcomes can result—an arms race, mutual disarmament, or an arms lead for one or the other side. The model supposes that both sides would most like to gain an advantage, would least like to be left behind, but would prefer mutual disarmament to an arms race.

The dilemma is as follows. Given our supposition that both players know what the game is, the leaders of Andersland may very well reason:

> I am not sure what Jedesland will do, but there are two possibilities, and either way, we do better by arming. If they disarm, arming gives us our best outcome ("4" instead of "3"). If they arm, then by doing the same, we at least avoid the worst outcome. Come to think of it, I can see that their leaders face the same problem. We must therefore expect them to arm too.

The argument from Jedesland's point of view is of course analogous. For each, arming represents a so-called *dominant strategy:* one that pays off whatever the other does. So it surely seems rational for each to arm, and to expect the other to do

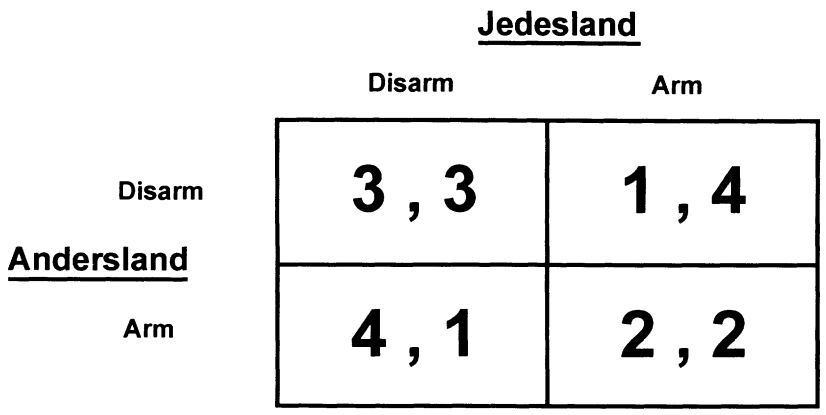

FIG. 3 Prisoners' dilemma game as model of an arms race.

the same. Unfortunately, both players then do worse (2,2) than they would if they both acted irrationally (3,3).

It is no surprise that Chicken and Prisoners' Dilemma are the most famous of all game models, both having acquired a huge literature of their own. Though obviously very simple in some senses (only two players, two strategies each, no conditional moves, and so on[1]), they seem to pose genuine dilemmas for rational choice. The response to such dilemmas, however, varies in theory as well as in practice. This reflects differing views about the purposes of analysis.

Purpose and Interpretation of Game Models

Among those who work with game models, there are some sharp divisions of opinion as to their main purpose. Recognizing that rational choice can be problematic need not rule out construction of normative theory. (Binmore, 1987, for example, elegantly makes the case for doing so.) Others argue that such an approach has led to a paucity of real applications, at least outside econometrics where heroic assumptions about rationality and quantifiability of values sit less uneasily than in IR. Thus, Schelling comments in the preface to the 1980 edition of *The Strategy of Conflict:*

> In putting these essays together . . . I had hoped to establish an interdisciplinary field . . . I wanted to show that some elementary theory could be useful not only to formal theorists, but also to people concerned with practical problems. I hoped too, and I think now mistakenly, that the theory of games might be directed toward application in these several fields. With notable exceptions like Howard Raiffa, Martin Shubik and Nigel Howard, game theorists have tended to stay instead at the mathematical frontier. The field that I had hoped would become established has

[1] Both are also *symmetric:* swapping over the players would not change the structure of the game. The widespread use of such examples occasionally leads to the misconception that games must be so. This is not true. Most games are asymmetric with respect to preferences, strategies, or both. As a consequence, the players may be in quite different positions of power.

> continued to develop, but not explosively, and without acquiring a name of its own (pp. v-vi).

My own sympathies lie with Schelling's lament. Arguably, many recent developments "at the mathematical frontier" could in principle promote the reorientation for which he had hoped. However, most theorists have a rather different agenda. Two linked issues can be discerned here: one to do with research strategy, the other with the aim of modelling problems of choice.

Many game theorists, stressing the importance of deductive theory, see their prime task as that of using the axioms of rational choice to construct *generic* models of particular types of situations. They will be more interested, for example, in clarifying the notion of deterrence and deducing results as to its stability under certain conditions than in trying to model an individual case directly. This top down approach obviously contrasts with attempts to apply the theory on a case-by-case basis, in which one has to make assumptions about specific actors' preferences and available strategies at a particular time.

Ideally, the two approaches should meet in the middle, but this rarely occurs. Both have strengths and pitfalls. Both face problems of empirical testability: generic models may be *too* generic to be confronted with real data, while case-based modelling can all too easily become a search for *post hoc* assumptions allowing one to get the "right" answer from the model. Since deductive theory is Game Theory's strong suit, there is little point in not using it where possible. Otherwise one is left with nothing more than a metaphor for describing situations—and an unattractive one at that. What is most important is that generic and specific models should not be confused. In some cases, it is not made clear which interpretation is intended. Understandably enough, those doing the first sort of analysis often mention specific cases in passing to illustrate the sort of issue being addressed. Equally understandably, critics see this as analysis of the specific case, carried out without the necessary empirical research and guilty of gross over-simplification. In making a general point about commitment, for example, the vicissitudes of particular decision makers may be of little concern. It matters little whether one calls the players Kennedy and Khrushchev, the United States and the Soviet Union, Andersland and Jedesland, or A and B. It also does not matter whether strategies are formulated in more than indicative terms such as "stand firm." Yet, in drawing practical conclusions about specific cases, these things matter a good deal. (See, for example, Reynolds' 1989 criticism of Zagare.)

Game theorists have had substantial successes in building elegant generic models and continue to extract further insights from them, even on topics on which one would have expected there to be little left worth saying. Recent examples include Klein and O'Flaherty's (1993) paper on the nature of threats and promises, Nicholson (1990) *inter alia* on deterrence and uncertainty, and Morrow (1989) and Powell (1989) on the dynamics of crises. The last is of particular interest in deducing, from plausible assumptions, results that contradict received wisdom. For example, there are circumstances in which a (rational!) challenger may be more likely to escalate the crisis when faced with a more resolute defender, and vice versa. Similarly, a challenger with a greater stake in the status quo need not therefore be less likely to contest it. Such general results are significant. Nevertheless, a theory which has nothing to say about individual and idiosyncratic cases would be a disappointment to most, and the rest of this review will presume some interest in such applications.

The second issue at stake is what one might hope to gain from the analysis of games. The initial aim was to find *uniquely rational* solutions to problems of choice.

Much of mainstream Game Theory has continued in this direction. For example, it is argued that dominant strategies *should* always be chosen unless one can make binding agreements to the contrary. The paradox in Prisoners' Dilemma is only apparent (Shubik, 1970). Thus, multiple equilibria can be dealt with either by refining the equilibrium concept itself, or by allowing players to randomize their choices. Uncertainty can always be allowed for through judicious use of subjective probabilities. Ultimately, there must be an answer. An alternative view is that models are better seen, not as ways of producing prescriptive solutions, but as ways of clarifying the structure of the situation, complete with whatever dilemmas and difficulties it may hold. The aim, then, is to help one understand the decisions faced by the actors, the interactions among their choices, and the tactics (threats, promises, bargaining, communication) they may employ to further their aims. This does not mean throwing rationality away, but it makes its application less ambitious. Recognizing that there may be good reasons for the players to arrive at various outcomes, analysis can explicate which results are supported by which lines of reasoning. This interpretation is adopted in the rest of this review. Though this approach side-steps some of the attacks made on Game Theory, the question remains whether game models can provide a fruitful framework for understanding. To this issue we now turn.

Criticisms and Responses

Forty-odd years' worth of controversy has yielded a rich literature critiquing, and defending, Game Theory. (See, for example, Green, 1966; Martin, 1978; Rapoport, 1964.) In essence, critics of Game Theory typically allege that its models are oversimplified, that the theory is inherently repellent, or that the whole notion of rational choice is unrealistic or morally unsound. Let us consider each group of charges in turn.

Limitations of Simple Models

Simple models have their attractions. It is difficult not to recognize real situations that are somewhat akin to Chicken or Prisoners' Dilemma. However, that is quite different from implying that, the Cuban Missile Crisis for example, was *nothing but* a game of Chicken—even one played for appallingly high stakes. If one wishes to analyze specific conflicts, the descriptive limitations of simple game models are well-enough rehearsed. Leaving aside the issue of whether the underlying assumptions—for example, those regarding preferences—are plausible, their format cannot encompass a good deal of what is typically known to be important. These missing dimensions can be summarized under four headings:

- *Differing perceptions:* Key participants may be acting according to very different definitions of the situation.
- *Dynamics:* The sequence in which decisions are made may be important, as may changes (in a wider sense) in the relevant players, their preferences, available options, and perceptions.
- *Combinatorial complexity:* Many possible specific actions may be open to each side, and these actions may or may not be independent of each other. There may also be many players.
- *Linked issues:* Actors may be simultaneously involved in many issues, within and between alliances, nations, governments, committees, and so on.

Criticisms based on the models offered in the 1960s often seem to assume that these complexities cannot be addressed. As will be seen, extensions to the model go some way to doing so. First, however, we briefly consider arguments that charge Game Theory with fundamental philosophical flaws rather than just with over-simplification.

Three Reasons to Dislike Game Theory

To many people Game Theory can appear highly quantitative. Some critics appear to have an inbuilt aversion to anything containing numbers. (Note that this need not be the case. An early and persistent critic was P. M. S. Blackett (1962), one of the founding fathers of Operational Research.) However, the use of mathematical structure need not mean quantifying everything in sight. In practice, many models avoid the need for much measurement. Preference orders can be used rather than utility scales. Often, results do not even depend on having complete orderings, but only on broad assumptions about what players like or dislike. Most theorists want their theories to apply as widely as possible. Some also go out of their way to make their work accessible to non-mathematicians. But, unfortunately, much of the specialized work in Game Theory remains impenetrable to the non-mathematician.

Secondly, just as numbers can repel, so can words. To describe an important problem as a "game" can be deeply offensive. Protestations that this is just a matter of wording often fall on deaf ears—and the theory does draw on *some* analogies with parlor games. Arguably, its terminology can contribute to hiding unpleasant realities behind anodyne words. "Game" also conjures up winners and losers, even though little of the theory is concerned with zero-sum cases, and much interesting work has been done on cooperation (notably Axelrod, 1984; Powell, 1991; and Snidal, 1991; *inter alia*). Nevertheless, some text-books still pay disproportionate attention to zero-sum games, and cover illustrations prettily superimposing military hardware onto chessboards are unlikely to attract the peace-loving.

Thirdly, the theory's frequent association with both free-market economics and with hawkish versions of deterrence can make it unpalatable to those on the political left. As a matter of history, many pioneers of Game Theory were deeply affected by the politics of the Cold War, which is hardly surprising given the times. (Poundstone provides a highly readable account, especially of Von Neumann's own role as a "cold warrior.") However, the record is not one-sided. Rapoport, for example, combined distinguished contributions to Game Theory and gaming with lively attacks on the Cold War mind-set (Rapoport, 1965, 1960). Even Schelling's early work on deterrence—though frequently read as a justification for toughness—retains a lively sense of the dilemmas and dangers of brinkmanship. More recent work has usefully contributed to debates about the security dilemma, and about the possible effects of misperception. As for economics, much of Game Theory does sit squarely within a free market paradigm. But game models can be used to undercut overly optimistic assertions about the market's "invisible hand." As Prisoners' Dilemma shows, individual rationality can produce an outcome disadvantageous to all. In this general vein, Hardin's *Tragedy of the Commons* (1968) has spawned a vigorous literature, much of it arguing the need for viable collective solutions (Lash and Urry, 1985). Game-based models can—if one so chooses—provide a useful medium for debating the balance between central planning and local autonomy. In summary, those unimpressed with deterrence theory and/or free market economics need not be hostile to Game Theory as such. Nevertheless, they will probably take the approach in directions rather different from the mainstream of current theory.

On Rationality, Selfishness, and Morality

Game Theory's presumption that players (particularly, in IR, states) are to be modelled as *rational actors* has attracted much criticism, and this issue warrants slightly lengthier discussion. Rationality is often taken to imply selfishness, and rational choice theory interpreted as offering an immoral (or at least amoral) view of life. One problem here is that there are many different uses of the term "rational." There is instrumental (means/end) rationality, procedural rationality, consistency of aims or beliefs, rationality of argument, and so on—each of which comes in different variants (Hargreaves Heap, et al., 1992; Nicholson, 1992; Selten, 1988). Some concepts of rationality are essential to some versions of Game Theory but not to others. For example, some forms of analysis do not even require preferences to be transitive (Howard, 1986).

On the face of it, Game Theory starts with a very strong notion of rationality. Zero-sum theory characterizes optimal strategies that players should, as a matter of compelling mathematical proof, follow in order to maximize their utilities. Players are assumed to have complete information about the game, and sufficient powers of calculation to analyze it fully. However, several of the more recent developments in Game Theory abandon the assumption of complete information, while that about reasoning power is not as demanding as it looks. If we take games to be subjectively defined, an actor's limited powers of reasoning can be modelled by keeping the perceived game simple. What *is* always central is instrumental rationality, relative to whatever game is perceived. Players are presumed to have preferences for at least some outcomes, and to choose accordingly. But contrary to a still-popular misconception, these preferences need not be selfish. Players may have many reasons for preferring some states of the world over others. Their wants need not be for material things at all; they just have to be definable. In particular, players' preferences may depend on benefits or disbenefits to others (so-called social factors). For example, some writers have contrasted an "institutional liberal" view of international politics, in which states attempt to maximize their individual benefit, with a "realist" view in which they also value relative advantage over others (Grieco, 1988). The specifics of this relative gain argument are controversial (Grieco, et al., 1993), but the point is that Game Theory allows one to apply either assumption and to track its consequences. Similarly, our previous arms race model assumes that each side wants to get ahead but regrets the expense of the race. If players are more motivated by goodwill (or conversely, actually enjoy being armed to the teeth), the result is another game with different preferences. The game format itself is no more or less applicable. (For some alternative models to Prisoners' Dilemma, see Wagner, 1983.) Indeed, this realization is essential to appreciate the *paradoxes* of rational choice, which can, for example, be constructed for altruistic players, not just to entrap the selfish.

Nevertheless, Game Theory's very lack of any prior presumption about motivation can rekindle the charge of amorality, in that players can be as benevolent or as homicidal as we like. In this sense, morality has to come from outside the theory. However, game models can help to clarify the consequences of different moral schema. For example, it can throw light on the Kantian notion that moral policies are those that are beneficial when generalized (Colman, 1982). Similarly, Harsanyi (1975) uses game-theoretic arguments to criticize aspects of Rawls' theory of justice. Explicit application of Game Theory to moral issues is still rather rare, though its logic is relevant to any ethical system concerned with the interaction of decisions.

A final charge, linked with that of amorality, is that of ignoring needs. Game Theory, it is argued, pays attention only to actors' wants or interests as expressed in preferences. But conflicts of interest are relatively superficial. Deep-seated

conflicts are about needs which can "neither be traded or suppressed" (Burton, 1985). Outside the IR context, neglect of needs has been attacked by some psychologists, notably Smail (1987), and even the occasional economist (Braybrooke, 1977; Scitovski, 1992). Though preference can be defined widely, it seems unsatisfactory merely to treat all needs, even those of which actors may be unaware, as contributors to preferences. It is then difficult to see how they can enter into game-based models, except perhaps as constraints on the range of acceptable outcomes. (Drama Theory, discussed below, may allow this incompatibility between needs and preferences to be reconciled to some extent.)

In summary, the game paradigm need not take players to be self-interested, materialistic, or cold-blooded. However, it does unashamedly stress the importance of deliberate choices (excepting evolutionary models in which strategies can be patterns of behavior adopted without conscious decision). Players are taken to be motivated by preferences for outcomes. They also recognize that they are in interaction with each other and plan accordingly. Comparative neglect of needs can be seen as an important limitation. This need not prevent one from using preference-based models, but it must be acknowledged that they only tell part of the story.

Enriching the Game Model

So far, we have presented a partial defense of Game Theory, while also noting various limitations. This section examines how far these limitations can be overcome by reasonably well-established extensions to the theory. It considers, in turn, the missing dimensions noted before: *differing perceptions, conflict dynamics, combinatorial complexity,* and *linked and multi-level games.* For the sake of continuity, it refers back to the Cuban Missile Crisis example at each point. While there is renewed debate about the crisis itself—ably reviewed by Scott and Smith (1994)—the generic points made below should be non-controversial. They are reflected in a wide variety of accounts, by participants and by both critics and admirers of U.S. policy (Abel, 1986; Allison, 1971; Bernstein, 1976; Kennedy, 1968). Various alternative game-theoretic *models* of the crisis have appeared, but we will not attempt to review these. The interested reader will find appropriate material in, for example, Brams (1985), Fraser and Hipel (1984), and Wagner (1989). This section concludes with some brief comments on modelling methodology.

Differing Perceptions

Of all the limitations of the basic game model, arguably the most fundamental is the assumption that all the players see the same "game." That is, they share the same view of what the issue is, who the relevant players are, what strategies they have, and what their preferences are. A cursory examination of real life suggests that this is seldom so, and many far-from-cursory studies of IR confirm the point (Axelrod, 1976; Boulding, 1959; de Rivera, 1968; Holsti, 1962; Jervis, 1976, 1970). This need not make "rational actor" models irrelevant, but it does suggest that rationality is a very subjective matter. In many cases perhaps decision makers do conceptualize situations in terms of players, strategies, and preferences, but see *different* games. Modelling the Cuban Missile Crisis as a game of Chicken has some plausibility, for example, simply because key figures on the U.S. side seemed to see it that way (as in Robert Kennedy's "We're eyeball to eyeball, and the other fellow just blinked"). But it does not follow that it was seen the same way by anyone else.

Not surprisingly, attempts to incorporate differing perceptions into Game Theory have a long history. For example, Luce and Adams (1957) considered some effects of misperceived payoffs, while examinations of deliberate deception are

many (Brams, 1977). From a mathematical point of view, Harsanyi's (1968) model of *games with incomplete information* has had a great impact on the subsequent development of Game Theory. His starting point is that players may be uncertain of each other's utilities. This is modelled by introducing a set of possible *types* of player. All the players know what this range of possibilities is, but they are unsure of each other's specific type. However, they can assign probabilities. For example, player A may believe that B must have preferences either as in Chicken (player type "1") or as in Prisoners' Dilemma (type 2). She then decides that there is a 75% chance that B is a type 1 player, a 25% chance that B is type 2. Misperceptions are modelled by allowing players to have different probability estimates of each other's type. This ingenious idea has been elaborated in a number of ways, and it has the attraction of dealing with uncertainty as well as misperception. However, the resulting analysis is complicated, and it rests on very strong assumptions about the structure and consistency of each player's beliefs. Both utilities and beliefs must be fully quantifiable and fully consistent for each player. Applications have tended to be in economics rather than IR.

An alternative approach is provided by *hypergame analysis* (Bennett, 1980, 1977). Whereas the previous model starts with the idea that there is a well-defined game going on, but the players may be uncertain what it is, hypergame analysis adopts the simpler and more radical *presumption that* actors construe the world differently. The starting-point is therefore a set of subjective games, one for each player. In principle, these games might have quite different players, strategies, and preferences. Nevertheless, they are connected in the sense that an action taken in one will have an impact on the others. The differences in perception may be localized, concerning preferences for particular outcomes or the feasibility of a particular strategy. For example, an alternative to the Prisoner's Dilemma arms race model is one in which each side really prefers mutual disarmament most of all, but believes *the other* to have Prisoner's Dilemma-type preferences (Bennett, 1977; Plous, 1987). Despite the players' real wish to disarm, each believes that the other prefers to cheat on any agreement. Cooperation may thus be prevented by mutual distrust. Strategic surprises in war provide obvious cases of different beliefs about the possible moves.

In other cases, however, the relevant issues may be construed in completely different terms. To take a slightly trite illustration, one actor might see the world in terms of competition between nation-states, another as a struggle between religions, and a third in terms of class interests. Each might have a perfectly consistent view of some issue in which they are all engaged. To understand their interaction, it is necessary to ask how actions translate from one world-view to another. For example, the Vietnam War seems to have been defined in terms of the battle against communism for some actors, but in terms of national unification for others. And each may have presumed that this was a shared definition.[2] Analysis needs to allow for the *possibility* of such wholesale differences. If the differences do not exist, it is easy enough to simplify the model.

Conflict Dynamics

That situations often change over time is clear enough. From a modelling point of view, we need to consider both the sequence of moves within a particular game and the wider point that the parameters of the conflict (for example, preferences)

[2]This is meant to imply that the whole conflict may be construed differently, not just that actors attach different weights to preference criteria (for example, one side presumes both see communism as more important than nationalism, the other, vice versa). The latter type of situation, of interest in its own right, requires only modest elaboration of the incomplete information model.

may change. *Within* a game, the traditional view was that all issues related to the sequencing of moves could be dealt with via the Von Neumann concept of strategy. Though strategies might have to be very complicated, any series of choices could in principle be telescoped into a single choice of strategy without changing the properties of the game in any way. This view has now been abandoned. In particular, it is acknowledged that strategies that are in equilibrium at the start of the game may cease to be so as moves within them are implemented. In other words, it might pay a player to revise her choice of strategy as the game unfolds. This realization, which challenges the original foundations of Game Theory, has led to various refinements of the equilibrium concept (Kreps & Wilson, 1982). Most important is the idea of a *Subgame-Perfect Nash Equilibrium (SPNE)*. As successive moves are made, the choices still available create a subgame. The strategies in an SPNE must remain in equilibrium in each successive subgame. This is a development of great importance, particularly for theorists in search of uniquely-rational solutions to games, who can argue that only subgame-perfect equilibria represent possible solutions because they are the only stable outcomes. More generally, it has led to renewed interest in looking at games move-by-move—that is, at the extensive-form game tree—rather than confining attention to the strategic matrix.

Developing the previous Cuban Missile Crisis model will illustrate both the analysis of a game tree and some of the previous points about differences in perception. Consider the following sequence of possible moves, starting from the point at which the United States (US) has become aware of the missile installations. Firstly, the Soviet Union (USSR) decides whether to complete the missile installations ("deploy missiles" for short) or withdraw. If the USSR deploys missiles, the US must decide whether or not to use force to remove the missiles—through invasion, air strike, or both ("attack Cuba"). If the US uses force, the USSR can choose to escalate the conflict further. The basic tree is shown in Figure 4 (a). Note that putting the moves in strict sequence is itself a simplification (for example, the US could in principle strike immediately), but it does give some sense of the dynamics of choice. This basic structure is similar to that used by Wagner (1991), though he develops it into a more elaborate model incorporating both uncertainty and learning about aims.

The preferences that guide the decisions of each side are attached to the outcomes at the end of the tree. Rather than suppose that the two sides agreed on what each other's preferences were, the model can be used to illustrate the popular contention that the Soviet leadership misunderstood U. S. preferences. That is, they believed that they could get away with placing missiles in Cuba without provoking a military response, thereby gaining strategic parity cheaply. The tree on the left of Figure 4(b) reflects this possible Soviet view. Preferences are shown, as in Figures 2 and 3, by ranking the four possible outcomes from 4 (best) to 1 (worst), with U. S. preferences on the left in each case. Thus, the USSR leadership would most like to keep the missiles with no military conflict. Failing that, it would rather withdraw straight away, but it would prefer to escalate if the US attacks Cuba. The US (the USSR leaders presume) wants the missiles withdrawn, preferably without the need for military action, but would rather accept them than face an escalated conflict.

Analysis is carried out by working backward from the end points of the tree. Here, such analysis would lead the USSR leadership to reason as follows:

> If we are placed in the position of having to respond to US military action (the bottom-left subgame), we would rather escalate than withdraw. So the US must expect that we would do so. The US would therefore be better off avoiding this branch of the tree by not attacking Cuba, since they prefer (2,4) to (1,2). Moving now to the top of the tree, we can choose given this prediction of the US response.

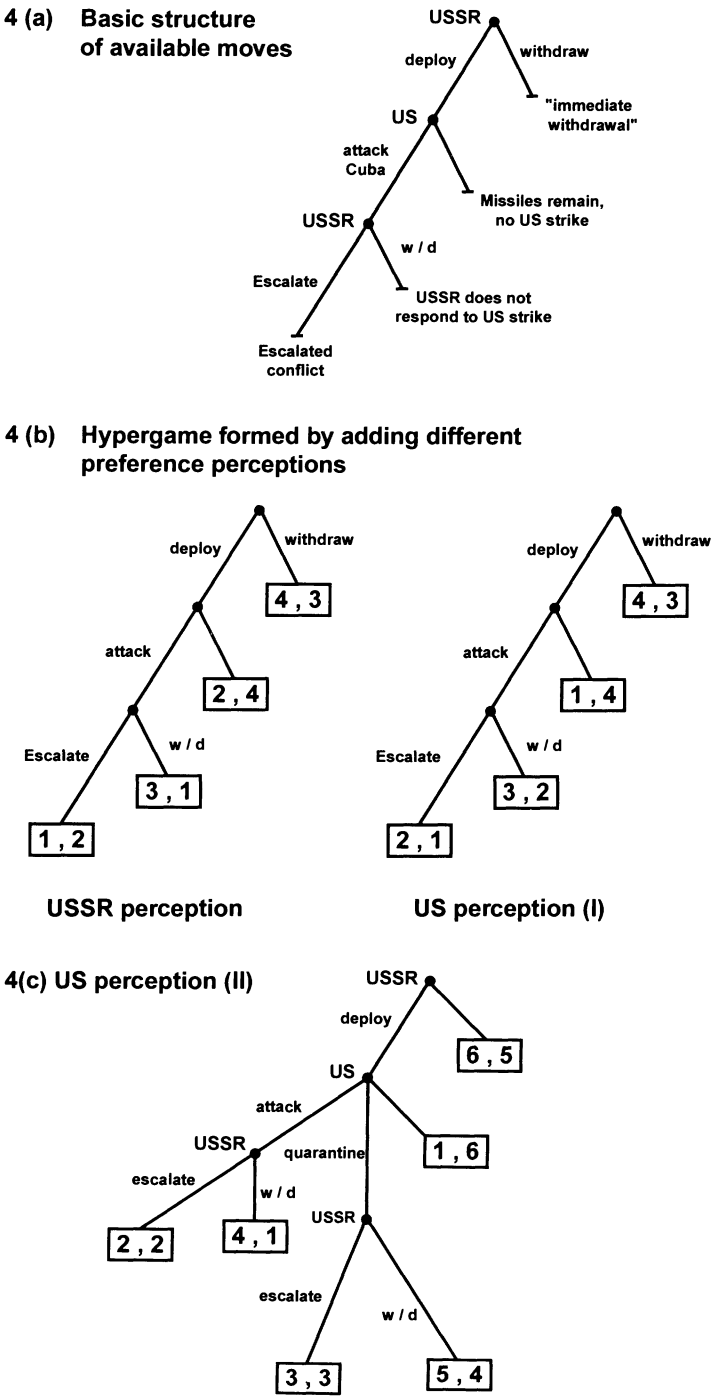

FIG. 4 Developing tree (extensive form) models for Cuban missile crisis.

> We prefer (2,4) to (4,3). So we keep the missiles, and the US will (eventually) back down.

History shows that the United States took a tougher line. We can interpret this in terms of a perceived game differing from the Soviets' in at least one of two ways: (a) the United States had a preference for escalation over accepting the missiles, and/or (b) the United States believed that the Soviets would prefer not to escalate in response to a strike against Cuba. It is plausible that at least the more hawkish elements on the U. S. side were inclined toward both views, in which case the perceived game tree would be as on the right of Figure 4(b). This results in a hypergame in which each side misconstrues the other's preferences (though in this simple model, the available moves are seen in similar terms). Given such a view, the US leadership could be expected to reason thus:

> If we attack the missile sites, the USSR could choose to escalate, but that would be against their interests. So we would prefer the anticipated result of attacking (3,2) to accepting the missiles (1,4). So the USSR would do better to withdraw straight away (4,3), which is what we want them to do.

Note that in this view, accepting the missiles is actually the *worst* outcome for the US. It would still be better to strike even if the USSR then escalated after all. The policy recommendation might be couched thus, "By making clear our willingness to strike, we can make the USSR withdraw. If they are irrational enough not to withdraw, we should strike at the missiles anyway and call their bluff on escalation."

Had these remained the views taken on each side (and both seem to match arguments made at some stages of the crisis), the potential for disaster is clear. Analyzing the hypergame as a whole, and taking account of each side's perception in turn, one sees an inexorable drive toward escalation, even though this is neither wanted nor predicted by either party. Fortunately, a more sophisticated view prevailed within the U. S. administration, a model of which is shown in Figure 4(c). Here, it is realized that the USSR would prefer to escalate if the US strikes at Cuba. However, a new alternative, "quarantine" has been invented for the US, against which it is believed that the Soviets would prefer withdrawal to escalation. (Otherwise preferences are as in Figure 4(b), allowing for the fact that there are six outcomes to order, rather than four.) History shows that this latter belief, at least, was justified. In this perceived game, the predicted outcome once the USSR had attempted deployment is quarantine and eventual withdrawal (5,4). Note also that had the USSR leadership perceived something like this game from the start, their best option would have been to avoid the crisis altogether (6,5). This matches the generally-accepted view that Khrushchev's position was weakened by his Cuban "adventure." The exact preference for escalation held by either side was never put to the ultimate test.

So far, we have looked only at the dynamics of moves within a particular game. The use of *multi-stage games* holds out the promise of more fully dynamic models that could allow, for example, for preferences to evolve as the situation unfolds. This and similar extensions to the game model are usefully reviewed by Thomas (1986). Models such as Morrow's (1991) on repeated crises combine two-stage analysis with attention both to incomplete information and what players may learn from each other's behavior. These extensions of the basic model consist of a series of linked games, with the outcome reached at each stage influencing which game the players face next. For example, certain outcomes of an initial trade negotiation game might lead to different varieties of trade war, others to various trading regimes, and so on. Each of these results is modelled, not as a fixed outcome, but as a consequential game. Analysis is a more complicated version of that for a game

tree, similarly working backward from the end points. Thus, one would analyze each of the final games to find the expected outcomes, then compare these to establish *which games the players would prefer to play.* The preceding game can then be analysed using both this preference input and any preferences for its immediate outcomes. The previous example could quite easily be extended in this way. In particular, the outcome rather vaguely entitled "escalation" could be better modelled as a game in its own right. In a hypergame, it is of course quite possible that players may have different views about what consequential games are possible. For example, a plausible interpretation of events leading up to the Falklands/Malvinas War is that the Argentinean leadership did not perceive (until too late) any possibility of ending up engaged in a full-scale war (Haig, 1984; Hastings and Jenkins, 1983). In general, multi-stage games can provide a useful compromise between treating different phases of a conflict in isolation and attempting to include everything in a single game of unmanageable complexity.

Combinatorial Complexity

Game trees and strategy matrices both become very complicated as the number of moves increases. Yet failure to deal with such complexity lends weight to the charge that game models oversimplify. Continuing with the Missile Crisis example, many possible actions were seriously considered. Even ignoring tactical detail, those considered on the U. S. side included a pre-emptive nuclear strike, an air strike against the missile sites, invasion of Cuba, diplomatic protests and negotiation, trading U. S. missiles in Turkey, and undertaking *not* to invade Cuba, along with various combinations of these options. And this is a case in which the starkness of the choices makes a simple model more than usually attractive! More significantly, the "quarantine" option actually chosen can be seen as a creative escape from the crude dichotomy between backing down and attacking. Similarly, although the crisis is unusually easy to treat as a two player case, other parties could arguably take significant actions. All this is consistent with Game Theory, but an approach reliant on a full definition of all available strategies, or a complete game tree, before any conclusions could be drawn would be impossibly cumbersome. Practical analysis requires a way of building up and analyzing models step-by-step, encouraging one to try out alternative assumptions.

One such approach is *Analysis of Options* (Howard, 1986, 1971; Radford, 1980), in which the actions open to each side are structured in terms of simple "yes/no" options presented in the form of a tableau. For example, Figure 5 shows a somewhat richer set of options for the Missile Crisis, with the number of parties considered increased from two to four (adding both the Cuban Government and US Allies). Outcomes (or scenarios) are represented by columns of 1s and 0s, denoting options chosen or not. Two are shown for illustration. Scenario A represents the situation at the height of the crisis. Reading *down the column,* this shows that the US *is* implementing the quarantine and diplomatic protest, while *not* invading, trading missiles, or giving a non-invasion assurance. Meanwhile the USSR has suspended shipments but is not removing the missiles, and so on down the column. Scenario B is one interpretation of the resolution of the crisis. Preferences for any player can be shown very easily through the order in which the scenario columns are displayed. By convention, those "more preferred" are placed to the left.

If there are n options in the model, each of which can be taken up or not, there will in principle be 2^n scenarios. However, not all of these will be *feasible.* Thus certain choices may be incompatible with each other, or only available in combination. This is important in helping to limit the complexity of analysis. In the tableau shown, a few of the option combinations that might be regarded as

	Scenario A	Scenario B	Infeasible combinations
US			
Quarantine Cuba	1	0	– – –
Diplomatic protest	1	0	– – –
Air strike	0	0	– – 0
Invade Cuba	0	0	– – 0
Trade Cuba and Turkey missiles	0	1	1 0 –
Give non-invasion assurance	0	1	– – –
USSR			
Suspend military shipments to Cuba	1	1	– – –
Remove missiles	0	1	– – –
Trade Cuba and Turkey missiles	0	1	0 1 –
Retaliate against US if Cuba attacked	–	–	– – 1
Cuba			
Request USSR to suspend operations	0	0	– – –
US Allies			
Continue to support US policy	1	1	– – –

*The three columns on the right show combinations that cannot be taken up ('infeasibles') Thus, the first pair indicate that a missile 'trade' must be agreed by both sides: the US cannot trade if the USSR does not, and vice versa. The third column means that if the US adopts neither air strike nor invasion, the USSR cannot 'retaliate'. In each case, the dashes in the rest of the column mean that the infeasibility holds whether or not these other options are taken up.

FIG. 5 Some options and scenarios for Cuban missile crisis.

infeasible are displayed on the right. For example, a missile trade requires both sides to implement, so scenarios with a 0 for one "trade" option and a 1 for the other are infeasible. (The dashes in the rest of the column mean that this is so regardless of which other choices are made.)

As shown schematically in Figure 6, analysis proceeds in stages:

- Starting from a particular scenario, we examine whether any participant has an *improvement* available, that is, can move unilaterally to a more highly-preferred scenario.
- Whenever there *is* an improvement (say for player P), we then identify any possible *sanctions* against it, that is, a response by another player that would leave P worse off than in the original scenario. If there are none, the improvement is "guaranteed." P cannot do worse by moving, regardless of how other parties respond.
- This process is repeated, by considering different possible improvements from different scenarios in turn. Where appropriate, one can also consider improvements and sanctions open to coalitions, as well as to individual parties.

In this way, conclusions about various scenarios are built up. Scenarios from which there is a guaranteed improvement are regarded as unstable, but all others are

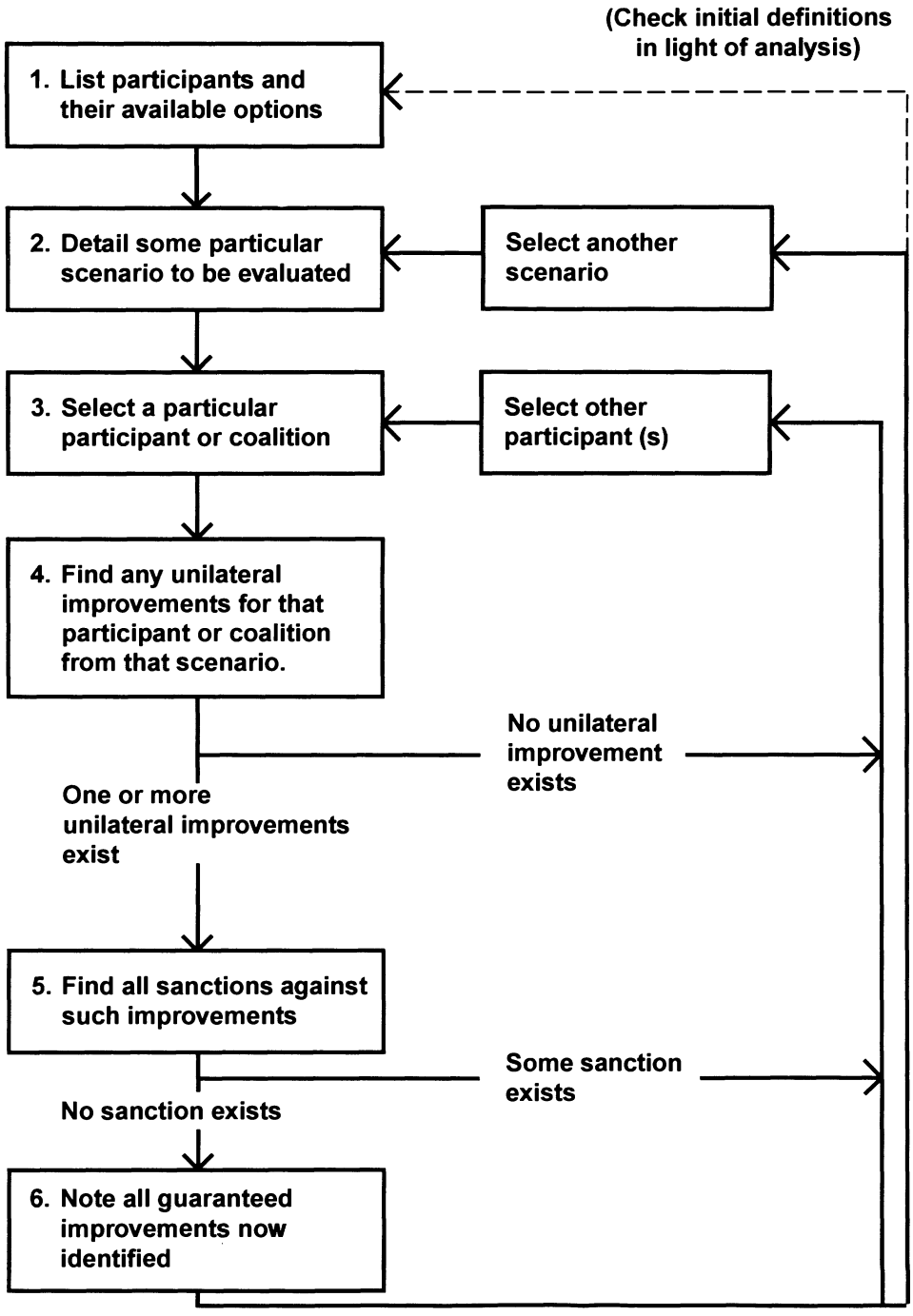

FIG. 6 Summary of basic analysis of options procedures.

considered as potentially stable because sanctions represent threats that *may* deter moves away from them. This leads to a very wide criterion of stability, including not only all Nash equilibria (those outcomes from which there are no unilateral improvements), but usually a good many other outcomes. The aim of the approach is to explore the reasons why particular scenarios may be viable, rather than to narrow down the predicted solutions as much as possible. Nevertheless, having established this wide set of stable scenarios, further steps can be added that reduce them somewhat. For example, one may go on to examine how credible particular sanctions are likely to be. This can either be done as part of the formal analysis, by presuming that credibility depends solely on the preferences of the threatener, or by adding informal judgments about particular actors.

This method provides a useful way of dealing with comparatively large numbers of options needing to be considered in combination—especially if the number of parties considered can be kept reasonably small. Some computerized versions of this model have appeared, and one is discussed later, using a more fully-developed example.

Linked Issues and Multi-level Games

The existence of linked issues is another commonly recognised fact of life. Thus, it is widely accepted that even the Missile Crisis can only properly be understood within a contextual web of other issues, both internal and external to each side. This not only multiplies the relevant players; it also complicates the decisions faced by each. The Kennedy administration was influenced by the pressures of bureaucratic politics, and by political imperatives such as the need not to be seen as "soft on communism." Linked international issues included the need to demonstrate support for a weak ally such as Turkey. For Khrushchev, the whole episode may have been an attempt to reconcile longstanding disagreements over national priorities.

As already noted, game models are often charged with ignoring all this by considering only monolithic, unitary national actors. Certainly, many models have been framed in such terms, and modellers have been influenced by the realist tradition in which nations *are* the relevant players (Jervis, 1988). But game theorists would be the first to acknowledge the existence of conflicts at many levels. In effect, every player is a game. Even some analyses ostensibly based on a unitary actor view—notably Snyder and Diesing's *Conflict Among Nations*—actually end up emphasizing internal factors. But although it is easy to acknowledge the linkages between issues, it is quite another matter to allow for them in a systematic way. Nevertheless, this is currently an area of much research activity. As regards internal and external influences on national diplomacy, a popular metaphor is that of leaders having to play simultaneous games at two separate tables. The linkage between the tables has been variously conceived, with stress laid on the domestic determinants of foreign policy or *vice versa*. These ideas are reviewed and extended in a highly-influential paper on "two level games" by Putnam (1988), followed up by the collection entitled *Double-Edged Diplomacy* by Evans, et al. (1993). What is new here is not the notion of linked games as such, but the investigation of specific hypotheses about the linkages, for example how the configuration of internal interest groups affects foreign policy decisions. As Putnam himself comments, the underlying model is (so far) a descriptive metaphor rather than a piece of analytical theory. Nevertheless, it may prove to be a very fruitful one in integrating empirical work. Other ways of structuring linked issues, both between and within organizations, are presented by Radford (1980) and by Bennett (1988) among others. These can serve as preliminary models delineating the relevant context of

analysis of specific issues. Examples of work at a more analytical level include models of the "guns versus butter" dilemma such as Powell's (1993). This bridges internal and international concerns by considering national leaders as having limited resources, so they must balance the pursuit of military strength with valued domestic programs.

Comments on Modelling Methodology

All the developments just sketched provide ways of producing more complex models. Yet simplicity has much to commend it. It is difficult to deny that something of the air of Chicken pervaded the Missile Crisis, and that this helps one interpret some of the decisions made. (The early stages of the Gulf crisis invited a similar interpretation of Saddam Hussein's behavior.) A richer interpretation may be given by elaborating the model, but there is no guarantee that a more complex model will perform better. The more data are required, the more difficult testing may become. In extreme cases, models may be so complex that their own behavior becomes inexplicable. From this perspective, the appropriateness of a simplification depends entirely on how well the resulting model works. The principle is well illustrated by Bueno de Mesquita's work on the outbreaks of wars (Bueno de Mesquita, 1981; Bueno de Mesquita and Lalman, 1992). Having eloquently attacked the notion of treating nations as unitary actors, he argues for an alternative, but equally radical, simplification: treating decisions to go to war as if they were made by individual key leaders maximizing expected utility. Taken literally, such models may be unrealistic. But if they nevertheless work well empirically, this is an important result.

Returning to the missing dimensions of simple games, we thus do not suggest that all complexities should be catered to in every model. Instead, the need is for a flexible framework within which different forms of complexity can be tested, so that one can clarify which are helpful in each case. This view of analysis is far removed from that of following a recipe: "specify all the players, enter all the options and preferences seen by each side, then analyze the system to find the answer." Models are supposed to simplify, and one can seldom predict which simplifications will prove most appropriate. Choices must be made as modelling proceeds. The important point is to know what *can* be included. This is all the more important when models are being used in a practical setting.

Avenues of Further Development

The additions outlined so far enrich the modelling repertoire, but there are arguments for more radical innovation. This more speculative section discusses three possibilities. Two of these approaches take one well beyond the original scope of Game Theory, while still making use of some of its deductive structure. Both approaches also stress the use of empirical theory. In this respect, it is important to note that Game Theory and its derivatives have little substantive content of their own. They provide a mechanism for deriving conclusions from assumptions about players, strategies, and preferences, but they offer no guidance as to what assumptions might reasonably be tried out—for example, what sort of strategies are likely to be considered in various circumstances, or what factors affect preferences. The two approaches discussed below, by contrast, start to address such issues. The final area of development lies in the use of computer software to assist formal modelling and analysis. We conclude with an example of applied research bringing together several of these ideas.

Drama Theory

In the sense used here, Drama Theory is a modelling approach integrating extended versions of Game Theory with research both on the role of emotion and on the strategic uses of argument (Howard, et al., 1992). Clearly, a change in terminology is involved. Developments within Game Theory already take one further and further from the everyday sense of the word "game." Given its unattractiveness, the case for change is strong. While "drama" carries some negative connotations, these are less marked.[3] To refer to the drama unfolding in Bosnia, for example, is surely more acceptable than calling it a "game," unless with deliberate irony.

However, more is proposed than just a change of metaphor. To the previous developments from Game Theory are added specific hypotheses about the dynamics of conflict, especially the origins and role of emotions and preference change. It is suggested that these are often triggered by the *paradoxes of rationality* within the situation perceived by the actors. The key is the need to employ unwilling threats and promises, whose implementation would actually run contrary to one's own preferences (like the threat to match toughness with toughness in Chicken or the promise to repay cooperation in Prisoners' Dilemma). Mainstream Game Theory simply dismisses such threats or promises as incredible. Using solution concepts such as the Subgame-Perfect Nash Equilibrium automatically has this effect. Drama Theory, by contrast, regards the need to make irrational (that is, contrary-to-preference) intentions credible as a centrally important source of both emotion and preference change. Its main features (Bennett and Howard, 1994; Howard, 1994) are as follows:

- A structure of parties (now termed *characters*), options, scenarios and preferences now represents a particular "frame" seen by a given character at a particular point. As in Hypergame Theory, the frames seen by the various characters may be quite different.
- The drama unfolds as a series of *episodes* in which some of the characters interact. As in a multi-stage game, their actions will influence not only the outcome of each episode, but also which episodes are encountered subsequently.
- Within an episode, each character adopts a particular *position,* trying to persuade the others to accept a particular scenario. However, characters typically have to use unwilling threats and/or promises to bolster their positions.
- The characters attempt to act rationally within their respective frames. However, at the *climax* of the episode, the paradoxes of rational choice create specific forms of emotion and preference change. They also generate new options, attempts to form new coalitions, etc. These change the frame. The characters now take actions which move them into another episode.
- Resolution of the drama occurs when the characters reach a common position that does not rely on unwilling threats or promises. This position is stable both rationally (in terms of choice within the current frame) *and emotionally* (in that there are no further paradoxes of choice to trigger changes in frame).

These ideas can be elaborated in various ways. Just as one can have linked games, for example, so dramas often exhibit distinct but interlinked sub-plots. However, the main feature of the model is its attempt to combine the analysis of choice within temporarily fixed frames with that of transformations from one frame to

[3]Varoufakis (1991) provides an interesting prior discussion of switching metaphors from game to drama, as well as a critique of "pure" rational choice theory.

another. That there can be change is hardly a new observation. What is new is Drama Theory's suggestions as to how specific forms of change are triggered. Basic hypotheses are:

- That relying on an unwilling threat or promise to support one's position will create pressure to change one's preferences, tending to make it willing.
- That such transformations are accompanied by negative or positive emotions toward the other side. Re-framing of the situation then tends to rationalize (make consistent) one's own actions and preferences, both to oneself and to others.
- That deceit is always an alternative to genuine change. Emotion change can be faked; argumentation may be insincere. However, faking is not cost-free and does not in general remove the need to examine the logic of genuine change.

In the short term, emotion can be seen as causing one to act irrationally—that is, against one's underlying preferences (as in Richard Nixon's "madman theory of deterrence"). In the longer term, underlying preferences may themselves change. One begins to hate the enemy enough to prefer destruction to defeat. Similarly, the need to make unwilling promises can generate positive emotion and eventual preference change in favor of altruism.[4] Though preferences can also change for exogenous reasons, this strategic source of change at least warrants attention and empirical test.

While more work remains to be done, existing evidence establishes a *prima facie* case for these hypotheses. For example, Frank (1988) assembles a wide range of evidence supporting the contention that emotional deviation from rational self-interest is not only widespread but beneficial. Not only can the capacity to signal emotions be advantageous, the benefit is not wholly negated by others developing the ability to fake. This is important for Drama Theory, since its contentions are dependent on deceit not being universally cost-free and effective. There is also direct evidence (much of it summarized by Cialdini, 1993) that people frequently alter their behavior, at considerable material cost, for no better reason than to maintain consistency with previously-stated opinions. The very issuing of a threat thus tends to produce preference change, making one more willing to implement the threat. Such subconscious pressure will also tend to be reinforced by conscious reasoning. An important means of establishing credibility is through *supporting argument* (Elster, 1991). For example, one may embed the message in arguments about one's own longer-term reputation. More subtly, a threat may be rephrased as a warning, pointing to a danger that all wish to avoid. It is scarcely plausible that constructing a convincing case will not affect one's own beliefs. *Publicly-stated* arguments carry further costs of reversal. As Elster notes, even the "threat rephrased as a warning" can become self-fulfilling because other actors take it at face value. In short, talk is often *not* cheap.

In using the term "drama," one is pushing onto an already-crowded stage. Some

[4]This interpretation takes preferences to be anchored in a system of *values*, rather than being merely behavioral. Following Fisher and Urey (1991), along with many others, preferences are seen as lying behind positions, and values behind preferences. The more deeply anchored one's preferences, the greater the emotion associated with change. Deciding to fight will be more traumatic for the committed pacifist, not just because of the magnitude of the preference change, but because the original preference was deeply-rooted. A geopolitical strategist, by contrast, might undergo the same change in preference merely by revising some balance-of-power calculations. If psychological *needs* for self-identity, etc., are considered as the most fundamental level of values, the need for resolution of deep-seated conflict, in Burton's (1985) sense, to be dramatic is clear. For needs cannot be changed without massive emotional, as well as intellectual, upheaval. The change may be rationalized afterwards but, unlike changes in surface preferences, it leaves trauma that cannot be rationalised away.

existing uses of the metaphor (Goffman, 1959) place less emphasis on choice, stressing the influence of predetermined scripts. Though this different usage may risk confusion, there are also opportunities to make new connections. For example, just as Game Theory can be used to analyze parlor games, one can fruitfully analyze scripted dramas, treating the characters as if they were real. Indeed, dramas of lasting interest generally have something to say about the human condition. This links with a growing body of Hermeneutical analysis of myths and legends as archetypes of real-life decisions (Alker, 1987, *inter alia*). Drama Theory may thus serve as a bridge between analyses often seen as belonging within quite different paradigms. The interest in how decisions are framed (and how frames change) also suggests links with research on framing from a decision-theoretic viewpoint (Faber, 1990; Tversky and Kahneman, 1979).

Knowledge-Based Systems

Drama Theory starts with a formal model and adds substantive theory to provide dynamics. A complementary approach would start by systematizing empirical knowledge about a conflict and then using it predictively. Advances in computing Artificial Intelligence can, it is argued, greatly aid the detection of patterns across events. This is presently a flourishing field (Carbonell, 1978; Cimbala, 1987; Hudson, 1991; Mefford, 1987). One can envisage a hierarchy of rules embedded in a computer system to guide the modelling of any given case, rather than relying on purely formal analysis. Such Expert Systems have achieved success in other complex fields, and the attractions in IR are considerable. For example, Schrodt argues that good short-term predictions can be routinely made using expert (or even semi-expert) knowledge of foreign policy, with "a level of intersubjectively validated accuracy, at least at the level of meteorology and quite possibly at the level of some fields of biology" (1987:91). Furthermore, such predictive success far surpasses that attainable using "rational actor" models, a claim explored more fully in Schrodt (1985). Rational actor models thus find themselves under attack from two sides: from those who have never liked formal models, and from those who regard them as *passe* in the light of new technology.

If human beings are indeed quite successful predictors, it makes sense to try to elicit the rules of inference they use, particularly if the rules can be systematized and laid open to critical scrutiny and improvement. If the result could then be encoded into software, the individual user could in principle have the wisdom of many experts available at her fingertips. This, in a nutshell, is the case for an Expert System approach. However, a number of objections can be raised against the case in its more ambitious versions. Expert opinions differ as to how much reliable predictive theory there really is about international conflict. Even where hypotheses have been extensively studied, controversies about evidential support remain. A case in point is research on the effects of crisis on decision makers, in which recent empirical work undermines much accepted wisdom. (Compare Brescher, 1993, with Milburn, 1972.) This suggests that IR is far from having a knowledge base as reliable as that for areas (such as medical diagnosis) in which Expert Systems have been most successful. This is now acknowledged by many working in the Expert Systems field. As Taber (1992:890) comments:

> Unlike experts in many other domains, U. S. foreign policy makers do not draw on a clear, consensual body of expert knowledge. Even in the 1950's, at the height of the Cold War, there was more than one way of thinking about the world. In general, multiple agents of different worldviews compete to produce policy outputs. These competing impulses arise from a set of quite different representations

> of the world. In fact, agents of different worldviews reason differently about the same input information and come to different conclusions.

Like others including Carbonell (1978), Taber responds by constructing systems which incorporate these competing worldviews, mirroring the shift within the rational choice paradigm to considering multiple perceived games.

Given this acknowledgement of the role of belief structures, it may not be too optimistic to suppose that useful *rules of thumb,* as distinct from laws of nature, can be formulated about the conduct of international politics. Useful findings arguably exist, spread across different disciplines. One might hope to establish a hierarchy of rules—some very general, some restricted to particular types of decision—or forms of intervention, such as third-party mediation. Though the result would formally resemble the knowledge base of an Expert System, there is a strong case for attempting something less ambitious. Rather than giving the user answers (and the reasons for them), one might concentrate on raising questions (and giving reasons why they may be important). Two arguments favor such a *Knowledge-Based Questioning System* (Bennett, 1991). First, the relative unreliability of the knowledge base would matter less if the aim were to provoke thought rather than provide answers. Second, raising questions may in any case be more useful. If humans are indeed naturally good at using patterns in past events to make predictions, then this may be something with which they need relatively little help. Studies of decision failures suggest that we are often less good at questioning the patterns we perceive. Alternative interpretations are frequently overlooked until too late. A helpful system would thus suggest alternatives, while leaving final selection to the human user. Avoiding the *term* "Expert System" might also help prevent users from having too-high expectations, just as many researchers already avoid the much-hyped phrase "Artificial Intelligence."

What role would this leave for formal modelling? It is possible to see formal analysis and substantive theory (pattern-matching) as symbiotic, rather than as rivals (Mefford, 1987). Firstly, the concepts of Game Theory can provide a common language within which rules of thumb could be expressed and applied. For example, Jervis (1988) uses Prisoners' Dilemma as a vehicle to bring together a range of propositions on topics such as the sources of preference and of preference changes. The work on two-level games discussed earlier similarly seeks to systematize empirical work, as does some of that on Drama Theory. Secondly, formal analysis can throw up some genuine surprises, for example about the stability of certain outcomes. Interactive decisions can be counter-intuitive enough for straightforward pattern-matching to be misleading.

Developments in Software

Developments in computer software are of great import both for formal analysis and Expert System approaches. The latter lie beyond the scope of this review, so we confine attention to software supporting game- or drama-based analyses. Building up and analyzing a tableau of options, or tree of moves, can be quite time-consuming, so the computer opens up the possibility of using more complex models. However, the previous arguments against pursuing complexity for its own sake still apply. So the more important benefit of new technology may be to allow a more flexible methodology, within which models incorporating different assumptions, or different forms of complexity can be rapidly built up, modified, and compared.

Recent software developments have improved the tools available to support such a methodology. An example is the INTERACT package (Bennett, et al., 1984).

Others include Fraser and Hipel's (1984) DecisionMaker, and the latest versions of Howard's (1989) CONAN. These can be used in various ways: for research, for teaching, to support experimental studies, and for decision support. The last use, decision support, demands a system that can be used as a vehicle for working *with* people engaged in the conflict, building and analyzing models on the spot (rather than going away and doing some analysis). With this in mind, INTERACT allows one to display and work on any part of the model at any time, with plenty of opportunities to add commentary on actors, options, scenarios, etc. Running in a windowed environment, it shows the relevant elements of the model on-screen, together with the available commands. This allows one to manipulate the model easily, while providing a constant visual guide to what is happening. The specific models used adapt existing pen and paper methods, though in the longer term other possibilities will be opened up by more advanced computer graphics. Experience so far shows that even the existing technology opens up quite new ways of using models in practice, as will be illustrated in the next and final section.

Modelling in Practice: An Example

We will now illustrate aspects of game-theoretic methods, such as Analysis of Options, and some use of Drama Theory, as well as the use of software such as INTERACT. Though one could take an example involving decision support, collaborative and field research with those engaged in an issue raises many of the same points. For illustration, we therefore take some current work (started in late 1993) on the conflict in Northern Ireland. This conflict has claimed over 3,000 lives during the last 25 years. Highlights of the eventful few months of late 1993 and early 1994 included the Shankhill bombing and Greysteel massacre,[5] the Hume/Adams peace initiative, the disclosure that the British Government had been engaged in secret indirect discussions with the Provisional IRA (Irish Republican Army), and the Downing Street Declaration establishing a joint policy position for the British and Irish Governments. Subsequent events, of course, have included the IRA and Loyalist ceasefires. Research on this new situation continues and should be reported in due course.

The work started as a piece of collaborative research with a clergyman interested in conflict resolution, who had established contacts with all the main political parties in Northern Ireland.[6] The initial aim was to use game/drama modelling to structure his knowledge of the situation, and to see what conclusions could be drawn. Using INTERACT, we first built up a structured picture of parties and issues. The software allows one to define issues within parties, to add, delete, or rename issues and parties, and to add commentary in plain text at will. Our working model focused on a four-sided constitutional issue involving the Catholic and Protestant Communities and the two National Governments. Internal issues included the

[5] These two events followed an all too familiar pattern. Spurred on by a series of attacks on Catholics in West Belfast, a local brigade of the Provisional IRA attempted to wipe out the leadership of one of the Loyalist paramilitary groups. Their bomb exploded prematurely, killing one of their own number and several bystanders, while sparing the intended targets. Loyalist paramilitaries then carried out revenge attacks with the explicit aim of killing Catholics. They succeeded in doing so by mowing down several people (including one Protestant) at the Greysteel bar/club.

[6] For those unfamiliar with the situation, these are:

- *Sinn Fein,* often described as the political wing of the IRA, and currently led by Gerry Adams.
- The *Social Democratic and Labour Party* (SDLP), the voice of constitutional republicanism and the main rival to Sinn Fein for the allegiance of the Catholic community in Northern Ireland.
- The *Ulster Unionist Party* (UUP), the more moderate of the two main Unionist parties (that is, those seeking to maintain Northern Ireland's union with the rest of the UK).
- The *Democratic Unionist Party* (DUP), personified by the more strident figure of the Reverend Ian Paisley. The DUP and UUP are rivals for the "leadership" of the Protestant community.

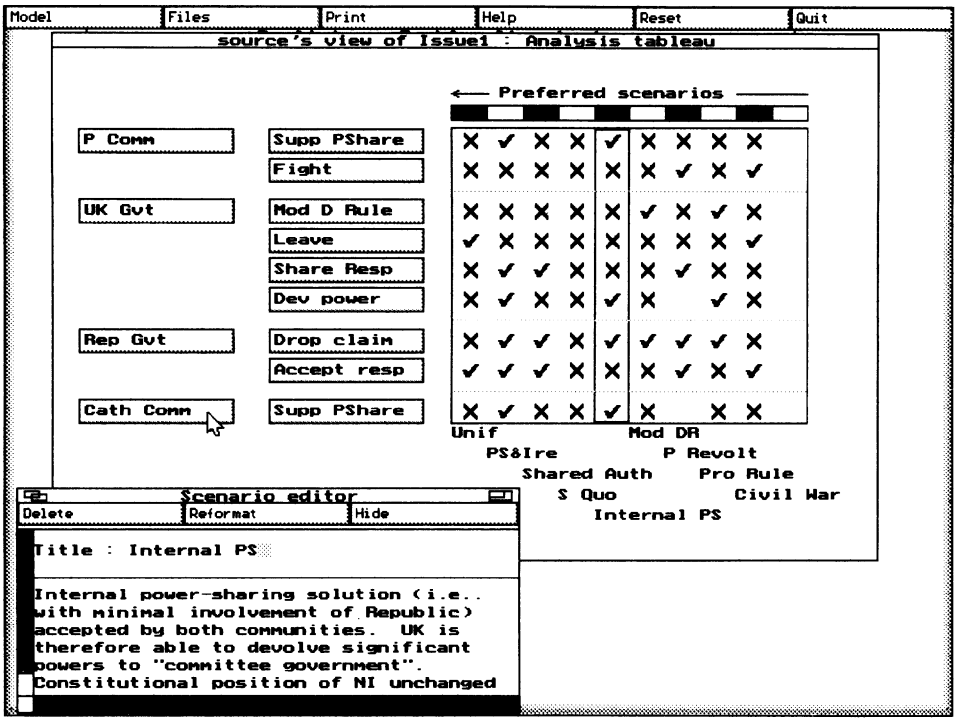

FIG. 7 Tableau for Northern Ireland analysis, showing typical INTERACT screen. Scenarios are shown in preference order for the highlighted actor, in this case the Catholic Community, with most preferred to the left of the tableau. Also shown is a typical "details editor" containing commentary on one of the scenarios. Brief explanation of some others: "Shared Auth" means that the two governments agree to govern jointly; "PS & Ire" means a Protestant/Catholic power-sharing assembly with the Irish Government given some executive powers; "Pro Rule" means that power is devolved to a Protestant-dominated assembly (as pre-1970).

rivalries between the political parties for the support of their respective communities, and that between the British Government and the opposition parties at Westminster. Prior to any formal analysis, this process provides a useful way of clarifying assumptions and building up a structured set of points for future reference. For example, questioning whether one could treat a religious community as an "actor" prompted clear argument about the nature of perceived religious identity, added in outline to the model.

The next stage is to choose a particular issue (and hence "level" of actor) for further analysis. We chose to concentrate on the constitutional issue. Modelling then follows the Analysis of Options format outlined above, in which binary options are drawn up in a tableau. In line with hypergame analysis, however, a different tableau can be used to represent each side's perception of options and preferences. Scenarios can be entered manually, or generated by the software. Figure 7 shows a typical screen, with a selection of scenarios placed in a presumed preference order for the Catholic Community. As in the Analysis of Options shown before, scenarios are read down the columns. However ticks and crosses are used instead of 1s and 0s to show options taken up or not. Thus the first column represents the unification scenario, in which the Protestant Community does not share power or fight, the UK government does not modify direct rule (from Westminster) but does

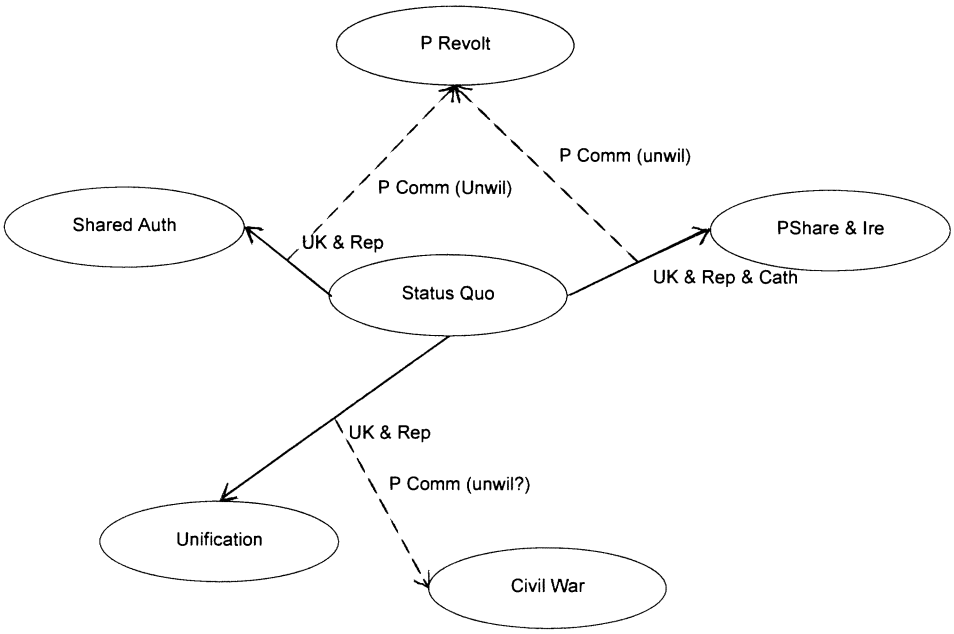

FIG. 8 Output of Northern Ireland analysis: a strategic map around the status quo.
From the status quo (as of 1993), the two governments in coalition have improvements to "Shared Authority" or "Unification." A coalition of both governments and the Catholic Community have an improvement to "P Share & Ire." All are shown as solid lines. The broken lines indicate sanctions available to the Protestant Community by choosing to "fight." The Protestant preference order used in most versions of the model makes these sanctions "unwilling," but they are not therefore dismissed as incredible.

leave, and so on down the column. The scenario editor window shows a portion of commentary added to one of the other scenarios, which can be called up and modified at any time. Each side's preferences can be put in by simply rearranging scenarios in the tableau (using the mouse), or can be generated from assumptions about whether particular options are liked or disliked.

The software will carry out analysis to find which scenarios are stable. It will identify improvements, list sanctions against them, and indicate whether these are willing or unwilling, providing the link from the so-far-static game model to Drama Theory. Following Howard (1987), the pattern of improvements and sanctions around any scenario is shown in the form of a *strategic map*. For example, Figure 8 shows a map generated around the then current status quo. Each scenario is now shown as an ellipse, with improvements and sanctions represented by solid and hatched arrows. Thus, for example, the UK and Republic Governments together can move to "Shared Authority," but the Protestant Community has the sanction of moving to "P (Protestant) Revolt" by choosing to fight.

Not surprisingly, the models indicated the existence of some unwilling (as well as willing) threats and promises. Examining these seemed to be useful in making sense of the behavior and tactics of the parties concerned. A notable point was the frequency with which the Protestant Community appeared to rely on unwilling threats (such as the sanctions appearing in Figure 8) to maintain its position. Exhibit 1 shows a few of the conclusions drawn at this stage. These are reproduced essentially verbatim to show how results stemming from the model can form a

EXHIBIT 1
Extract from working notes, initial analysis of Northern Ireland conflict

Analysis of Model 1 starting from current status quo:

No scenario appears preferred by all sides to the current status quo, which is thus (not surprisingly) stable in the sense that any improvement for some parties is subject to sanctions by others. A move to internal **power-sharing** requires Protestant participation. Our assumption is that they would not prefer this (at least, not in a form that would be accepted as genuine by the Catholic Community). Though all other parties prefer this, there is nothing in the model to induce the Protestants to move. However, one way forward would be to try to think of ways either to make the status quo less desirable, or power sharing more so. (Subsequent models look at situations in which the status quo has become less tenable.)

Unification and **Shared Authority** are (also) preferred by all except the Protestant Community. Both would be infeasible unless the Republic reforms. We suppose that the Republic would nevertheless prefer such a scenario. If the two Governments try to move in this direction, the Protestant Community has a sanction—choose to fight. The question then is whether this is "willing:" would the Protestant Community actually *prefer* to fight?

The answer can obviously vary between individuals. There may well be "ultras" who would be spoiling for a fight. But for many, it would be an unwilling threat, as in the version of Model 1 shown. However an unwilling threat can still be made credible:

(a) By expressing it not as an intent, but as a warning of something outside one's own control. For example, hard-liners on one's own side will run wild—("I abhor violence myself, but . . ."). Here, the credibility of this point is reinforced by existing action—by the UVF being prepared to attack police, etc., not just the Catholic Community.

(b) Through strategic use of emotion (which does not stop it being genuine, though it can sometimes be faked). In this respect, it is interesting just how angry Unionist politicians get at the merest hint of alternatives to British status (for example, reaction to the recently-leaked Labour Party options document). Such anger need not **just** be a knee-jerk reaction, or to please one's own constituents. It also serves as a small demonstration of just how angry one would get if *really* pushed hard.

(c) Through the longer-term pressures of trying to make an unwilling threat credible which tend to lead to preference change. Thus, the more that alternatives to British status are bandied about, the more the Protestant Community's preferences are likely to harden.

The last point, if valid, poses a conundrum. If the constitutional guarantee is seen as solid, the Protestant Community has no incentive to shift its position. If it is thrown into doubt, Protestants are under pressure to be less reasonable.

A British Government wanting to signal the necessity for the Protestants to accept change could try the same sort of tactics:

(a) By arguing that change will have to happen, for reasons unconnected with Government policy, such as demographics. (See Model 2.)

(b) By asserting that the Government *cannot* continue to underwrite the Guarantee ("regardless of what we would wish, we cannot continue to pay the price"). However, this smacks of weakness in the face of violence.

(c) By becoming publicly so exasperated as to appear not to care what happens, provided the mainland British are out of it. That is, the Government could try to convince the Unionists that (unless they are prepared to change) **abdication**—unilateral withdrawal—would be an improvement for the UK. Such a dramatic shift in preference is not inconceivable and would attract some popularity.

But it would be difficult for the British Government to make options (b) or (c) credible after so many years of paying the price.

Regarding unification itself, there is a further important point. Even if the British Government were undeterred by the Protestant threat, a move to unification requires the participation of the Republic. Despite its declared aspiration, the Republic is not that dissatisfied with the status quo and has everything to lose if unification went wrong. *It seems probable that the threat of fighting would deter the Republic Government.* Why should the Republic risk cooperating in its own destruction? Similar arguments apply to shared authority, though with less force.

framework for discussing both how existing behavior can be interpreted, and what might happen to change the original assumptions. This leads one to develop further models, in an iterative process.

The results to this point have proved interesting enough to lead to an invitation to take the process further. It was agreed to take the model on the road by visiting party contacts in Belfast, with the model on a notebook computer. We were able to discuss the model with politicians from each of the four main parties—in settings ranging from City Hall to distinctly battle-scarred premises in West Belfast—and to modify it as we spoke. The intention was both to improve the initial model and to find out how actors on the ground might respond to the modelling method. It would be improper to recount all that was said in detail, and in any case the research continues. However, some conclusions can be offered.

Firstly, portable computing power combined with flexible software does open up new possibilities for work in the field. It proved possible to explain the analysis (for example, how the tableau structured the issue). Most of those who saw it seemed engaged, particularly on seeing how the initial model could be modified on the spot. All seemed intrigued by the *idea* of using a computer model. The sense of novelty could be useful in gaining an entrée, though care is needed to avoid analysis becoming a gimmick and above all, any implied ability to conjure up a solution. These comments are not specific to the use of INTERACT. Other existing packages have their own strengths, and all will doubtless be superseded by further developments.

Secondly, there are advantages and disadvantages to starting with an existing model. Doing so runs the risk of opening with a view so unacceptable as to terminate discussion. In principle, building separate models from scratch would be better in allowing each side to reflect its own view, expressed in its own terms. Conversely, an existing model can give people something to react to, so provoking a dialectic. In this case, everyone seemed to see similar enough constitutional options for the initial model to be recognizable enough to discuss, though the alternatives (as well as preferences) were modified. The end result is thus a composite model explicitly containing the different perceptions expressed. (The most significant differences concerned the nature of Loyalist paramilitary violence.)

Thirdly, work in field research may have implications for the use of modelling in mediation. This has long been a topic of interest, both to modellers and practicing mediators (Mitchell and Webb, 1988). We were emphatically not attempting to mediate, or even act as unofficial go-betweens. Nevertheless, some parallels can be drawn. Various modes of mediation support have been proposed, from analysis done in plenary session with all participants, through backroom support for the mediator alone. In the middle are modes of working not too far from that described here, with investigative modelling of different views informing a mediator's efforts to construct an overall perspective. While well beyond what was attempted here, this appears a promising area for future research.

Concluding Comments

Starting from Game Theory in its most basic form, we have outlined many (but by no means all) of the recent developments from it. Taken in combination, conceptual and technical innovations hold out the promise of more effective methods for analyzing complex conflicts. While maintaining an emphasis on choice and decision, they also take one far beyond what is usually understood by rational choice theory. It might be helpful, therefore, to refer to a broader approach of *reflective choice theory.* There are many directions for further development of such an ap-

proach. Some will doubtless prove to be dead ends, but the chances are that some will lead to new insights into the interactive decisions that make up the tapestry of international politics.

References

ABEL, ELIE. (1986) *Missile Crisis.* Philadelphia: Lippincott.

ALKER, HAYWARD R., JR. (1987) Fairy Tales, Tragedies and World Histories. *Behaviormetrica* 21:1–28.

ALKER, HAYWARD R., JR., ROGER HURWITZ, AND KAREN ROTHKIN. (1993) "Fairy Tales Can Come True: Narrative Constructions of Prisoner's Dilemma Game Play." Paper presented at the Annual Meeting of the American Economic Association. Annahein, California. January.

ALLISON, GRAHAM T. (1971) *Essence of Decision: Explaining the Cuban Missile Crisis.* Boston: Little, Brown and Company.

AXELROD, ROBERT. (1984) *The Evolution of Cooperation.* New York: Basic Books.

AXELROD, ROBERT, ed. (1976) *The Structure of Decision.* Princeton, N. J.: Princeton University Press.

BENNETT, PETER G. (1991) Modelling Complex Conflicts: Formalism or Expertise? *Review of International Studies* 17:349–364.

BENNETT, PETER G. (1988) "Hypergames as an Aid to Mediation." In *New Approaches to International Mediation,* edited by Chris Mitchell and Keith Webb. New York: Greenwood Press.

BENNETT, PETER G. (1980) Developing a Model of Conflict. *Futures* 12:489–507.

BENNETT, PETER G. (1977) Towards a Theory of Hypergames. *Omega* 5:749–751.

BENNETT, PETER G., AND NIGEL HOWARD. (1994) "Rationality, Emotion and Preference Change: Drama-Theoretic Models of Choice." 13th European OR Conference. Strathclyde University, Glasgow. July.

BENNETT, PETER G., ANDREW TAIT, AND KEIRAN MACDONAGH. (1984) INTERACT: Developing Software for Interactive Decisions. *Group Decision and Negotiation* 3:351–372.

BERNSTEIN, BARTON J. (1976) We Almost Went to War. *Bulletin of the Atomic Scientists* 32:13–21.

BINMORE, KEN G. (1987) *The Economics of Bargaining.* Oxford: Blackwell.

BLACKETT, P. M. S. (1962) "A Critique of Some Contemporary Defence Thinking." *Encounter.* Reprinted in *Studies of War.* London: Oliver and Boyd.

BOULDING, KENNETH E. (1959) National Images and International Systems. *Journal of Conflict Resolution* 3:120–131.

BRAMS, STEVEN. (1985) *Superpower Games.* New Haven: Yale University Press.

BRAMS, STEVEN. (1977) Deception in 2 x 2 Games. *Journal of Peace Science* 2:171–203.

BRAYBROOKE, DAVID. (1977) "Reconciling Attention to Needs with Attention to Preferences." Third International Conference on Metagame Theory. York University, Toronto, July.

BRECHER, MICHAEL. (1993) *Crises in World Politics: Theory and Reality.* Oxford: Pergamon Press.

BUENO DE MESQUITA, BRUCE. (1981) *The War Trap.* New Haven: Yale University Press.

BUENO DE MESQUITA, BRUCE, AND DAVID LALMAN. (1992) *War and Reason.* New Haven: Yale University Press.

BURTON, JOHN. (1985) "World Society and Human Needs." In *International Relations: a Handbook of Current Theory,* edited by M. Light, and A. Groom. London: Pinter Press.

CARBONELL, JAIME G., JR. (1978) POLITICS: Automated Ideological Reasoning. *Cognitive Science* 2:27–51.

CIALDINI, ROBERT B. (1993) *Influence: Science and Practice.* 3rd Edition. New York: Harper Collins.

CIMBALA, S. J., ed. (1987) *Artificial Intelligence and National Security.* New York: Lexington Books.

COHEN, M. D., J. G. March, AND J. P. OLSEN. (1972) "A Garbage Can Model of Organizational Choice. *Administrative Sciences Quarterly* 17:1–25.

COLMAN, ANDREW. (1982) *Game Theory and Experimental Games.* Oxford: Pergamon Press.

ELSTER, JON. (1991) "Strategic Uses of Argument." Stanford Conference on Barriers to Conflict Resolution. February. (Proceedings in Press.)

ELSTER, JON. (1979) *Ulysses and the Sirens.* Cambridge: Cambridge University Press.

EVANS, PETER B., HAROLD K. JACOBSON, AND ROBERT D. PUTNAM. (1993) *Double-Edged Diplomacy: International Bargaining and Domestic Politics.* Berkeley: University of California Press.

FABER, JAN. (1990) On Bounded Rationality and the Framing of Decisions in International Relations: Towards a Dynamic Network Model of World Politics. *Journal of Peace Research* 27(3):307–319.

FISHER, ROGER A., AND WILLIAM UREY. (1991) *Getting to Yes.* 2nd Edition. London: Business Books, Ltd.

FRANK, ROBERT H. (1988) *Passions within Reason: the Strategic Role of the Emotions.* New York: W. W. Norton.

FRASER, NIALL M., AND KEITH W. HIPEL. (1984) *Conflict Analysis: Models and Resolutions.* New York: Elsevier.

GOFFMAN, E. (1959) *The Presentation of Self in Everyday Life.* Garden City, N.Y.: Doubleday.

GREEN, P. (1966) *Deadly Logic.* Columbus: Ohio State University Press.

GRIECO, JOSEPH. (1988) Anarchy and the Limits of Cooperation. *International Organization* 42: 485–507.

GRIECO, JOSEPH, ROBERT POWELL, AND DUNCAN SNIDAL. (1993) The Relative Gains Problem for International Cooperation. *American Political Science Review* 87:729–743.

HAIG, ALEXANDER M. (1984) *Caveat: Realism, Reagan and Foreign Policy.* London: Weidenfeld.

HARDIN, GARRETT. (1968) The Tragedy of the Commons. *Science* 168:1243–1248.

HARGREAVES, HEAP, S., M. HOLLIS, B. LYONS, R. SUGDEN, AND A. WEALE. (1992) *The Theory of Choice: A Critical Guide.* Oxford: Blackwell.

HARSANYI, JOHN C. (1975) Can the Maximin Principle Serve as a Basis for Morality? A Critique of John Rawls' Theory. *American Political Science Review* 69:594–606.

HARSANYI, JOHN C. (1968) Games with Incomplete Information Played by "Bayesian" Players. *Management Science (A)* 159:320, 486.

HASTINGS, MAX, AND S. JENKINS. (1983) *The Battle for the Falklands.* London: Joseph Press.

HOLSTI, OLE. (1962) The Belief System and National Images. *Journal of Conflict Resolution* 16: 244–252.

HOWARD, NIGEL. (1994) Drama Theory and its Relationship to Game Theory. *Group Decision and Negotiation* 3:187–206, 207–253.

HOWARD, NIGEL. (1989) The Manager as Politician and General: the Metagame Approach to Analysing Cooperation and Conflict, and The CONAN Play. In *Rational Analysis for a Problematic World,* edited by Jonathan Rosenhead. New York: Wiley.

HOWARD, NIGEL. (1987) The present and future of metagame analysis. *European Journal of Operational Research 32:1–25.*

HOWARD, NIGEL. (1986) The Usefulness of Metagame Analysis. *Journal of the Operational Research Society* 37:430–432.

HOWARD, NIGEL. (1971) *Paradoxes of Rationality.* Cambridge, Mass.: MIT Press.

HOWARD, NIGEL, PETER G. BENNETT, JIM W. BRYANT, AND MORRIS BRADLEY. (1992) Manifesto for a Theory of Drama and Irrational Choice. *Journal of the Operational Research Society* 44:99–103, and *Systems Practice* 6:429–434.

HUDSON, VALERIE. (1991) *Artificial Intelligence and International Politics.* Boulder: Westview.

HUXHAM, CHRIS S., AND D. MACDONALD. (1992) Introducing Collaborative Advantage. *Management Decision* 30(3):50–56.

JERVIS, ROBERT. (1988) Realism, Game Theory and Cooperation. *World Politics* 4r0:317–349.

JERVIS, ROBERT. (1976) *Perception and Misperception in International Politics.* Princeton, N. J.: Princeton University Press.

JERVIS, ROBERT. (1970) *The Logic of Images in International Relations.* Princeton, N. J.: Princeton University Press.

KENNEDY, ROBERT F. (1968) *Thirteen Days: A Memoir of the Cuban Missile Crisis.* New York: W. W. Norton.

KLEIN, DANIEL B., AND BRENDAN O'FLAHERTY. (1993) A Game-Theoretic Rendering of Promises and Threats. *Journal of Economic Behavior and Organization* 21:295–314.

KREPS, DAVID M., AND R. WILSON. (1982) Sequential Equilibria. *Econometrica* 50:863–894.

LASH, S., AND J. URRY. (1985) The New Marxism of Collective Action: a Critical Analysis. *Sociology* 18:33–50.

LAX, DAVID A., AND JAMES K. SEBENIUS. (1986) *The Manager as Negotiator.* New York: Free Press.

LUCE, R. DUNCAN, AND E. W. ADAMS. (1956) The Determination of Subjective Characteristic Functions in Games with Misperceived Payoff Functions. *Econometrica* 24:158–171.

MARTIN, B. (1987) The Selective Usefulness of Game Theory. *Social Studies of Science* 8:85–95.

MEFFORD, DWAIN. (1987) "Analogical Reasoning and the Definition of the Situation: Back to Snyder for Concepts and Forward to Artificial Intelligence for Method." In *New Directions in the Study of Foreign Policy,* edited by Charles F. Hermann, Charles W. Kegley, and James N. Rosenau. Boston: Allen & Unwin.

MILBURN, THOMAS W. (1972) "The Management of Crises." In *International Crises,* edited by Charles F. Hermann. New York: Free Press.

MITCHELL, CHRIS R., AND KEITH WEBB, eds. (1988) *New Approaches to International Mediation.* New York: Greenwood Press.

MORROW, JAMES D. (1989) "Bargaining in Repeated Crises: A Limited Information Model." In *Models of Strategic Choice in Politics,* edited by Peter Ordeshook. Ann Arbor: University of Michigan Press.

NICHOLSON, MICHAEL. (1992) *Rationality and the Analysis of International Conflict.* Cambridge: Cambridge University Press.

NICHOLSON, MICHAEL. (1990) Deterrence, Uncertainty and Cooperative Behavior. *International Interactions* 15:227–242.

ORDESHOOK, PETER C., ed. (1989) *Models of Strategic Choice in Politics.* Ann Arbor: University of Michigan Press.

ORDESHOOK, PETER C. (1986) *Game Theory and Political Theory: An Introduction.* Cambridge: Cambridge University Press.

ORDESHOOK, PETER C., EMERSON NIOU, AND GREGORY ROSE. (1989) *Balance of Power.* Cambridge: Cambridge University Press.

PLOUS, S. (1987) Perceptual Illusions and Military Realities. *Journal of Conflict Resolution* 31:5–33.

POUNDSTONE, WILLIAM. (1992) *Prisoner's Dilemma: John Von Neumann, Game Theory and the Puzzle of the Bomb.* Oxford: Oxford University Press.

POWELL, ROBERT. (1993) Guns, Butter and Anarchy. *American Political Science Review* 87:115–131.

POWELL, ROBERT. (1991) Absolute and Relative Gains in International Relations Theory. *American Political Science Review* 85:1303–1320.

POWELL, ROBERT. (1989) "The Dynamics of Longer Brinkmanship Games." In *Models of Strategic Choice in Politics,* edited by Peter Ordeshook. Ann Arbor: University of Michigan Press.

PUTNAM, ROBERT D. (1988) "Diplomacy and Domestic Policy: The Logic of Two-Level Games." *International Organization* 42:427–460.

RADFORD, K. JIM. (1980) *Strategic Planning, an Analytical Approach.* Reston, Va.: Reston Publishing Co.

RAPOPORT, ANATOL. (1965) *Strategy and Conscience.* New York: Harper and Row.

RAPOPORT, ANATOL. (1960) *Fights, Games and Debates.* Ann Arbor: University of Michigan Press.

REYNOLDS, C. (1989) Review of F. C. Zagare's *The Dynamics of Deterrence. Review of International Studies* 15:67–74.

DE RIVERA, JOSEPH H. (1968) *The Psychological Dimension of Foreign Policy.* Columbus, Ohio: Charles E. Merrill.

SCHELLING, THOMAS C. (1960/1980) *The Strategy of Conflict.* First Edition/Revised Edition with a new preface. Cambridge, Mass.: Harvard University Press.

SCHELLING, THOMAS C. (1978) *Micromotives and Macrobehavior.* New York: W. W. Norton.

SCHRODT, PHILIP A. (1987) "Pattern Matching, Set Prediction and Foreign Policy Analysis." In *Artificial Intelligence and National Security,* edited by S. J. Cimbala. Lexington: Lexington Books.

SCHRODT, PHILIP A. (1985) "Adaptive Precedent-Based Logic and Rational Choice: A Comparison of Two Approaches to the Modeling of International Behaviour." In *Dynamic Models of International Conflict,* edited by Urs Luterbacher and Michael D. Ward. Boulder: Lynne Rienner Publishers, Inc.

SCITOVSKY, TIBOR. (1992) *The Joyless Economy.* 2nd Edition. Oxford: Oxford University Press.

SELTEN, REINHARD. (1988) *Models of Strategic Rationality.* Dordrecht, Netherlands: Kluwer Academic Publishers.

SHUBIK, MARTIN. (1984) *Game Theory in the Social Sciences: Concepts and Solutions.* Cambridge, Mass.: MIT Press.

SHUBIK, MARTIN. (1970) Game Theory, Behavior and the Paradox of the Prisoner's Dilemma: Three Solutions. *Journal of Conflict Resolution* 14:181–193.

SMAIL, DAVID. (1987) *Taking Care: an Alternative to Therapy.* London: J. M. Dent.

SCOTT, LEN, AND STEVE SMITH. (1994) Lessons of October: historians, political scientists, policy-makers and the Cuban Missile Crisis. *International Affairs* 70:659–684.

SNIDAL, DUNCAN. (1991) International Cooperation among Relative Gains Maximisers. *International Studies Quarterly* 35:387–402.

SNYDER, GLENN H., AND PAUL DIESING. (1977) *Conflict Among Nations.* Princeton, N. J.: Princeton University Press.

TABER, CHARLES S. (1992) POLI: An Expert System Model of US Foreign Policy Belief Systems. *American Political Science Review* 86:888–904.

THOMAS, LYN C. (1986) *Games: Theory and Applications.* Chichester, U. K.: Ellis Horwood.

TVERSKY, AMOS, AND DANIEL KAHNEMAN. (1979) "Rational Choice and the Framing of Decisions." In *Rational Choice: the Contrast Between Economics and Psychology,* edited by Robin M. Hogarth and Melvin W. Reder. Chicago: University of Chicago Press.

VON NEUMANN, JOHN, AND OSCAR MORGENSTERN. (1944) *Theory of Games and Economic Behavior.* Princeton, N. J.: Princeton University Press.

WAGNER, R. HARRISON. (1989) "Uncertainty, Rational Learning and Bargaining in the Cuban Missile

Crisis." In *Models of Strategic Choice in Politics,* edited by Peter Ordeshook. Ann Arbor: University of Michigan Press.

WAGNER, R. HARRISON. (1983) The Theory of Games and the Problem of International Cooperation. *American Political Science Review* 77:330–346.

ZAGARE, FRANK C., ed. (1990) Special Issue on "Modelling International Conflict." *International Interactions* 15(3–4).

ZAGARE, FRANK C. (1987) *The Dynamics of Deterrence.* Chicago: University of Chicago Press.