CHAPTER

11

Game Theory or the Theory of Interdependent Decisions

CHAPTER OUTLINE

INTRODUCTION: ORIGINS, DEVELOPMENT AND RELEVANCE

Even among approaches in IR espousing positivism and science, game theory was the most controversial. Detractors say it is more a branch of pure mathematics than of IR. But game theory in IR is not just reflection of a desire to flaunt higher mathematics in IR.[1] It is relevant to IR both contextually and substantively. At the height of the Cold War, when direct communication between the two superpowers was infeasible, and guessing the move of the other was important for either, game theory is supposed to have proved handy. Substantively, by and in itself, game theory has been said to possess special relevance for the theoretical study of IR because both seek to understand conflict and cooperation. As two commentators point out:

> A common assertion in the study of international relations is that the choices of actors are interdependent. This interdependence leads to a strategic reasoning which becomes quite complex…, even for simple interactions. Game theory provides us with a comprehensive toolbox that allows in-depth explorations of such interactions…It often makes intelligible processes that appear puzzling, this without attributing causality to factors such as incompetence, irresponsibility, or lack of concern of decision-makers. For instance, the well-known security dilemma in Realism can be illuminated by the study of a prisoner's dilemma. And the use of games helps us understand the conditions of application of the famous 'tying hands' principle in international negotiations.[2]

IR was attracted to game theory as a 'formal theory of strategic interaction', because it 'provides a tool for the mathematical analysis of situations where all the actors can influence the outcome and so must consider others' actions when deciding what to do'. It is from this perspective that concepts and issues such as 'balance of power, democratic peace, deterrence, ethnic conflict, hegemonic stability theory, power transition theory and reciprocity' are approached through game theory.[3] One enthusiast regards it as a branch of rational choice theory, which not only assumes that actors are rational and select the strategy that profits them most in terms of consequences of their behaviour but presupposes for the actors 'no special abilities to make commitments' and, hence, helps analyse anarchical situations. Besides, many international interactions not only involve a limited number of actors but often entail bilateral settings,[4] where game theory provides its best insights.

As for disciplinary origins, after an initial analysis of a duopoly by Antoine Cournot in 1838, mathematician Emile Borel's important suggestions in 1921, and John von Neumann's insightful 'theory of parlour games' in 1928, modern game theory came into its own with the masterpiece of mathematician John von Neumann and economist Oskar Morgenstern, entitled *Theory of Games and Economic Behavior* (1944).[5] A large amount of initial spadework was done during the Second World War at Princeton, alongside work on nuclear physics by theoretical physicists.

BASIC TERMS

In game theoretic parlance, a *game* signifies a social situation involving two or more individuals, who are called *players*, and about whom two basic assumptions are made. They are *rational* and *intelligent*, both in a technical sense: *rational* because he/she 'makes decisions consistently in pursuit of his/her objectives', which is maximizing 'the expected value of his/her own pay-off' and *intelligent* because 'he/she knows everything that we know about the game and he/she can make any inferences about the situation that we can make'. Both the assumptions presuppose a utility scale, captured in the 'expected utility maximization theorem', which is dependent on a '*sure-thing* or *substitution* axiom' that Myerson paraphrases in this way:

> If a decision-maker would prefer option 1 over option 2 when event A occurs, and he would prefer option 1 over option 2 when event A does not occur, then he should prefer option 1 over option 2 even before he learns whether event A will occur or not.[6]

Neumann and Morgenstern named their strategy 'the theory of games of strategy', which brings out the essential point that the games 'require the players to choose among strategies'. But game theory could also justifiably be called 'theory of interdependent decisions', since a game controllable by the decision of any one of the actors is trivial, not meriting analysis. Once the strategies are available and their possible outcomes for a player are indicated and listed, and the function taking him/her from one to the other is properly defined, the game is there.[7] The game models can function on the basis of a few principles: (a) the rationality of the players is 'transitive', that is, in similar circumstances they will always prefer 'A' to 'B', 'B' to 'C' and so on. (b) Even if the players are not aware in advance of the strategies of the opponent(s), they are aware of the benefits and losses that may accrue for them from the strategies pursued or pursuable by the opponent, because the 'utility' to be gained from each of the strategies in the matrix can be guessed from the value assigned to each of them ('imputation' in game theoretic parlance). (c) The gains or losses of each of the players can be easily measured by some mathematical methods and formulae. Since none of the actors possesses complete control over events, each of them needs to reckon with the others' possible actions. Where Actor A's most beneficial course of action depends on what Actor B decides to do, the reverse is equally true, and both appreciate this situation: A will seek both to anticipate and influence B's choices, knowing that B is seeking to do the same from the other end. So even when Players A and B never meet, their decisions would interact, and they will face an outcome determined by both of their choices, amidst chances of mutual threats, deceit, bluff and counterbluff. But this does not tell the whole tale, because conflict is seldom divorced from discernible common interests and interactive decisions ensuring joint gains from cooperation, 'collaborative advantage' or at least keeping conflict within limits, conceivable.[8]

MAJOR TYPES OF GAMES

Games are classifiable in terms of the number of players involved, the nature of interaction and the nature of distribution of benefits into: (a) two-person and *n*-person games with *n* greater than two; (b) cooperative (or coalitional) and non-cooperative (non-coalitional) games; and (c) zero-sum (or fixed-sum) and non-zero-sum (or variable-sum) games, and the resultant four types of games: *two-person zero-sum*, *two-person non-zero-sum* (or *variable-sum*), *n-person zero-sum* and *n-person variable-sum* games. In a zero-sum game, the sum of the utility of Player A (+1) and Player B (–1) is zero (as in a two-handed poker or duel). But in non-zero-sum games, the players can win or lose in varying amounts. Game theory differentiates those 'situations in which a decision-maker acts independently from all other decision-makers' from those in which 'multiple decision-makers act like a group'. In the former, called non-cooperative games, the players cannot enter into binding agreements to impose some action on one another, while in the latter, they can do so for joint randomizations or undertakings to play or not to play certain strategies. The distinction is rooted in John Nash's assertion that cooperative game theory 'is based on an analysis of the various coalitions which can be formed by the players of the game'.[9] Besides, games can be in normal form and extensive form. When a matrix depicts a strategic situation where choices have to be made simultaneously, in technical terms this is the normal/strategic form of the game. To grapple with sequential choices, game theory normally resorts to extensive-form games or game trees. However, an extensive-form game can be converted into a normal-form game in a way that will preserve the configuration of the extensive form, in which rows and columns stand for plans for how to play the extensive-form game from start to finish.[10] When in chess, two very inexperienced players are evaluating each position from the point that is reached in the course of the game, they are playing it in sequential or extensive form. But at a higher level, the players are able to visualize several moves in advance for at least a few possibilities and, theoretically speaking, each should be able to draw up a separate list of all the mixes of manoeuvres or strategies available for the next three moves. This will generate a fast 'ramifying tree of alternatives', curtailing the game to an advanced final and immutable decision by each player of his/her strategy that will be invariably followed in the game. For von Neumann said that only this conceptual curtailment of the game to normal form has rendered it fit for mathematical treatment.[11] As he further pointed out, a normal form game can be treated as a special type of extensive form game, because one can associate a normal form game with each type of extensive form game.[12]

Two-person Zero-sum Games

In two-person zero-sum games (henceforth, TPZSG), whose mathematical tractability has facilitated the emergence of a 'beautiful' theorem,[13] called 'minimax', making it prominent in game theory, what one player gains the other player loses, making the sum zero. Although in the interest of simplicity, the players' utilities are thought to be comparable, such an assumption is not logically urgent. Customarily, the game is visualized from the angle of the first player and his/her pay-off, the second player's pay-off being the negative of it. The game in normal form is constituted by a set of strategies for each player. If Player A has left three strategies (x_1, x_2 and x_3) and Player B has made available four strategies (y_1, y_2, y_3 and y_4), then for each intersection of strategies x_i and y_j, there is a pay-off a_{ij}. Forward captures the situation in an artificial rectangular array or matrix with apparently artificial examples (Figure 11.1).

Without any theoretical aid, the best strategy for either player is not obvious, and trial and error seems to be the only way to find it, where seeing Player A is allured by the high scores in the third row and plunges for Strategy x_3, Player B pounces for Strategy y_4 for assured pay-off 0, which is his/her best possible score in that row and so on. A more satisfactory way of setting about it may, according to Forward, be as follows:

1. Look at each row in turn to identify the *lowest* entry, which represents the worst that Player B can do in response to the strategy chosen by Player A.
2. Then mark out the highest of these identified entries, which is the *maximin* or the highest utility that Player A can assure for himself/herself by right selection of strategies, which in the contrived example is 2, the outcome ensuing from the combination of Strategies x_1 and x_3.
3. Moving through the corresponding routine from Player B's angle, identify the highest entry in each column, which is the worst for him/her and then the lowest of the marked-out entries, the best for him/her, that ism the *minimax*.

	y_1	y_2	y_3	y_4
x_1	8	5	2	3
x_2	–5	2	0	5
x_3	9	9	1	0

FIGURE 11.1 A Zero-sum Game Matrix, Where Player A's Pay-off Can Be Grasped from the Strategies of Both the Players, the Pay-off for Player B Being in the Negative

Source: Adapted from Nigel Forward, *The Field of Nations: An Account of Some New Approaches to International Relations* (New York, NY: Palgrave Macmillan, 1971/2016), 24–26.

4. If the two values are at all identical, as shown here, then for each player the correct play is choice of strategies bringing out this value, in the example, Strategies x_1 and x_3.

This point in the matrix is a *saddle point*, which would be evident if one considers the matrix as a mapped area in which the pay-off represents the height of each point above the sea level; lowest looking horizontally and highest looking vertically; and emerging only if the minimax and the maximin procedures converge at the same value point. This then can be conceived as the *value* of the game and the pair of strategies visualizing this point can be described as the *solution*. Since neither of the players can better his/her position above this point by changing his/her strategy, this is also called the *equilibrium point*.[14] This has been described by Luce and Raiffa in the following terms:

> There exists a number v (the value of the game), a pure or mixed strategy (maximin strategy) for the row player which guarantees him at least v, and a pure or mixed strategy (minimax strategy) for the column player that guarantees that the row player gets at most v. These strategies are in equilibrium, and any pair of strategies in equilibrium yield a maximin and minimax strategy for the row and the column player, respectively.

But since such a saddle point based, tidy solution does not always automatically emerge, a solution is created by devising the concept of a 'mixed strategy', which essentially is a 'probability mixture of the listed pure strategies'. Here, rather than endeavouring to maximize his/her pay-off, each seeks to maximize his/her expected pay-off. In games that have no saddle point, optimal mixed strategies get into equilibrium, because neither player can improve his/her expected pay-off till his/her opponent settles for a non-optimal strategy. So players receive no extra benefits and their opponents receive no extra hurts if they have advance knowledge that their opponents are plunging for their optimal strategy.[15]

Talking about TPZSG is not complete without a few words about bluffing, which helps a man with a weak hand to get the better of his opponent by bidding excessively high irrespective of a strong or weak hand, thereby maximally enhancing the opponent's uncertainty; bluffing in games can be treated as a randomized mixed strategy.[16]

Two-person Non-zero-sum Games

The moment one steps out of the tidy corner of TPZSGs, either by increasing the number of players or by making the outcome non-zero sum, the whole idea of a mathematically derived solution, whether through minimax theorem or through a combination of pure and mixed strategies, becomes fuzzy. In two-person non-zero-sum/non-variable-sum games (TPNZSG), gains and losses no more need to be equal. 'For such games (called non-strictly competitive game by Luce and Raiffa), it is impossible to choose the utility functions of the players so that they sum to zero.... Most economic, political and military conflicts of interest can be realistically abstracted into game form only if their non-strictly competitive nature is acknowledged'. But contrary to naïve expectations, this requirement of agreement seldom simplifies the game except in extreme, trivializing instances of perfect concurrence/agreement. Partial agreements may complicate the issue to such an extent that an 'elegant and cohesive theorem' like minimax in competitive games becomes a far cry, necessitating invocation of such non-theoretic caveats as 'bargaining psychologies of the individuals', 'interpersonal comparisons of utility' and so on. This excessive indeterminacy of TPNZSGs alongside sophisticated mathematical models for TPZSGs has proved challenging to economists, sociologists, psychologists and so on who have, however, hesitated in offering slapdash generalizations.

In games that are strictly competitive, it is unthinkable that players can derive any mutual benefit through cooperation in whatever form, though in non-strictly competitive games such mutual gain has always a chance, making ascertaining if the players are allowed to cooperate or not imperative. A cooperative game is one marked by full freedom of pre-play communication of the players for making joint binding agreements, while a non-cooperative game is one where no such pre-play communication is allowed. In the latter type, indicating the two players by 1 and 2, their particular strategy sets by A = $\{\alpha_1, \ldots, \alpha_m\}$ and B=$\{\beta_1, \ldots, \beta_n\}$, and the outcome linked to (α_i, β_i) by O_{ij}, and assuming that each player has choices among mixes of outcomes that lead to a linear utility function, they let a_{ij} express the utility of outcome O_{ij} for Player 1 and b_{ij} for Player 2; Luce and Raiffa got a table of this description (Figure 11.2).

Two of the best-known examples of non-strictly competitive TPNZSGs are the *battle of the sexes* (BOTS) and PD. In BOTS, the game may, according to Luce and Raiffa, look like as follows.

	β_1	β_2
a_1	(2, 1)	(–1, –1)
b_2	(–1, –1)	(1, 2)

Non-strictly Competitive TPNZSG: BOTS

Source: Adapted from R. Duncan Luce and Howard Raiffa, *Games and Decisions: Introduction and Critical Survey* (New York, NY: Dover Publications, 1957/1985), 104.

Here, a man (Player 1) and a woman (Player 2) have to choose independently from two choices in an evening's entertainment plans. He can go to a preferred prize fight (α_1 and β_1), she to a preferred ballet (α_2 and β_2) but going alone spoils the fun. The game

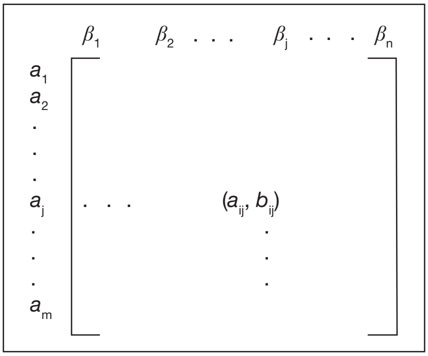

FIGURE 11.2 Two-person Non-zero-sum Non-cooperative Game

Source: Adapted from R. Duncan Luce and Howard Raiffa, *Games and Decisions: Introduction and Critical Survey* (New York, NY: Dover Publications, 1957/1985), 89.

does not contain any one of the characteristics of TPZSGs. If 1 declares that he is choosing and stubbornly sticking to α_1 and 2 believes his stubbornness enough, she has no choice but settling for β_1. The same thing happens if 2 does just as 1 did, because in such situations the player who reveals his/her strategy first with a touch of pre-commitment and with a reputation of inflexibility has an initial advantage. Even pre-play communication does not help, since although it is conceived to occur beyond the game structure of pay-off matrices, still in certain circumstances, it may occasion a fundamental change in one of the players' choice patterns and, thus, in the pay-off matrix. In such events, the 'space of strategies' may need to be enlarged, and the game may need to be further complicated to accommodate the pre-play negotiations.

When the players do not resort to any pre-play consultation and ponder their choices independently, that is, prefer the non-cooperative version of the game, the situation does not much improve. Space considerations prevent analysing in detail why both players separately decide to yield and land up with the (α_2 β_1) pair, consider optimizing the security level by choosing a mixed strategy or taking 'double cross' strategies (α_2 β_2), back off (since the mutual pay-off would be transformed into (–1, –1), reconsider playing 'safe' maximin strategies and then finally switch to 'double cross', again. Luce and Raiffa comment that the difficulty lies in the fact that 'the pair of maximin strategies ($x^{(0)}$, $y^{(0)}$) is not in equilibrium'.[17] Amidst this unending regress that breaks down all the comfortable features of the zero-sum case, von Neumann posited that once the husband and wife have established the negotiation set, the cooperative stage is no more amenable to mathematics. This 'goes for pure strategies and mixed strategies alike'.[18]

But others have argued that the main difficulty with games like BOTS is that they may have more than one equilibria, any one of which, if expected by both the players, may be a 'self-fulfilling prophecy'. In the present one, there are three sets of Nash equilibrium: (a) in both going to prize fight; (b) in both going to ballet; (c) in both playing in an uncoordinated and random manner, being unsure where the other would like to go, and getting the expected return from going either way. Being worse for both the players than either of the two foregoing equilibria, the third equilibrium is an inefficient equilibrium. To Myerson, in comprehending such games, the crucial question is what can motivate the players to expect each other to choose some specific equilibrium. He cites Schelling who argued that in games of multiple equilibria, anything that induces the players to concentrate on one particular equilibrium and make both of them expect and implement it like a self-fulfilling prophesy is called 'the focal point effect'. It may be socially caused, for example, in a society where wives have customarily truckled to the wishes of their husbands. So whether or not the couple feels impelled to conform to this tradition, it makes the (α_1 β_2) equilibrium more focal, and so more likely to be pursued.[19] In these circumstances, no calculus can go inside the negotiating set they have established.

A well-known example of the TPNZS-non-cooperative game is PD, which we will explicate while discussing the applicability and relevance of game theory to IR theory. For the present, let us move to *n*-person games, both zero sum and non-zero sum.

Word Help

Nash equilibrium—a Nash equilibrium comprises a list of strategies, one for each player, having the property that no player can unilaterally change his/her strategy and obtain a better outcome.

Mixed strategies—A game in normal form does not always contain a Nash equilibrium in which each player determinedly selects one of his/her strategies. The players may instead randomly select from among these pure strategies with certain probabilities. Randomizing one's own choice in this fashion is termed as mixed strategy. Nash showed in 1951 that any finite game in normal form finds an equilibrium if mixed strategies are allowed, an equilibrium being defined by a (probably mixed) strategy for each player where no player can gain on average by unilateral deviation. Average or expected pay-offs have to be considered because the outcome of the game may be random.

Source: Theodore L. Turocy and Bernhard von Stengel, 'Game Theory', CDAM Research Report LSE-CDAM-2001-09 (8 October 2001): 12 and 17.

N-person Zero-sum Games

It seemed intuitively reasonable to theorists including Rapoport that the two most critical concepts of TPZSG theory, namely mixed strategies and equilibrium points, could be applicable for games involving more than two players, and *n*-person zero sum games are not as irresolvable as they appear, since it is 'possible to view a two-person game as a special case of the *n*-person game (i.e., with $n = 2$)'.[20] Another commentator says, 'In two-person games, non-cooperation between the players corresponds to two coalitions—player 1 by himself and player 2 by himself—whereas cooperation means formation of the coalition of players 1 and 2 together'.[21] But others find 'qualitative differences between two-person and *n*-person cases, because only in *n*-person games may the players be motivated to split into alliances, blocs, cliques, factions, teams, gangs, caucuses, unions, syndicates or cartels whose collective interests diverge'.[22]

All these statements are differing perceptions of the fact that the *n*-person zero-sum game with $n > 2$ is always reducible to a sort of TPZSG simply by shifting main attention away from individual players to a coalition. Actions of coalition are different from those of individual players, first by the ability of members to maximize their collective pay-off, simply by coordinating their strategies, and second, by their ability to make side payments to each other. In consequence, although the allocation of pay-offs at the end of the game ('imputation' in game-theoretic parlance) will not necessarily accord with the outcomes or returns determined by the mix of strategies selected in terms of the pay-off function, the sum of the pay-offs would be the same. This is because of the presupposition of 'transferable utility' implicit in the idea of coalitions. Accordingly, von Neumann and Morgenstern focused on their theory of coalitions elaborately in their book to offer useful models for the study of economics, like oligopoly. They started from the premise, in sync with the rationality assumption of two-person games that the moment a coalition starts in an *n*-person game, all the other players will rush to build an opposing coalition, making it perfectly rational to design a two-person zero-sum game of Coalition A against the rest and go the whole process of a minimax theorem to work out the value of the coalition or the returns that Coalition A can assure itself of by employing a maximin strategy.

Of course, which coalition is likely to be formed in such situations and how the returns would be apportioned between the members remain open issues, even though von Neumann identified a few preconditions that would bring down the flurry of permissible solutions to 'manageable proportions'.[23] Luce and Raiffa too comment that a major impediment to 'developing a satisfactory theory of coalition formation is that in the present formalizations of a game no explicit provisions are made about communication and collusion among the players…Thus any theory of collusion, i.e. of coalition formation, has a distinct ad hoc flavor'.[24]

In the backdrop of the discouraging state of theory in this area, two creditable attempts to apply *n*-person theory to politics merit attention. The first is that of Shapley value. Circumventing the question of how it is derived from the value of the coalition by extension from the two-person game to an *n*-person one, I say that in one interpretation, the Shapley value is nothing but the '*a priori* value to each player of playing a game with a particular characteristic function and coalition structure'. In another, it is 'the *a priori* value that each player contributes to the grand coalition in a game with a particular characteristic function'. Viewed from each way, it betokens 'a rational way of dividing a pay-off among the players according to these *a priori* values'. If all the possible ways through which all possible coalitions could come into existence are registered and considered equally probable, then simple arithmetic could calculate the 'average expected advantage' that would accumulate for the coalition through the accession of a Player P. Since some members may have greater impact on the strength of the coalition than others, it is possible to consider the valuation of each player as 'an equi-probability mix of his/her influence relative' to the game, compared with that of other players.[25]

The second notable attempt of applying *n*-person theory to politics is credited to William Riker, who employed the logic of von Neumann and Morgenstern to political situations to derive his famous 'size principle' which predicts how minimum winning coalitions are formed. Standing in cooperative game theory, Riker viewed the formation of political coalitions as a fixed-sum bargaining game, where participants have to decide on the quota of valuable things, such as seats in the government. In Riker's own words, 'In social situations similar to n-person zero sum games, with side payments, participants create coalitions just as large as will ensure winning and no larger'. The size principle operates through two sub-rules: 'the strategic principle' and 'the disequilibrium principle'. Undersized parliamentary coalitions would be inclined to add new members, since the deprived majority would need both the resources and the incentive to topple the minority government (strategic principle). Contrarily, oversized coalitions may realize that amidst dwindling benefits of government membership, they need to throw out surplus members for greater dividends (disequilibrium principle). Despite the falling popularity of Riker's theory because of its assumptions of pay-offs emerging from coalition bargaining being zero sum, coalition members always securing a positive pay-off and uncomfortable assumptions of complete and perfect information, a hypothesis of him that has been tested with supportive results is: 'The greater the degree of imperfection or incompleteness of information, the larger will be the coalitions that coalition-makers seek to dorm and the more frequently, will winning coalitions actually formed be greater than the minimum size'. The model has other shortcomings, stemming from employment of the concept of winning in an unsophisticated manner; location in cooperative game theory; and the challenge from emergence of a considerable amount of recent scholarship from non-cooperative game models, paying closer attention to 'the incentives of individual parties' and elaborating each stage of the bargaining process.[26] But still, Riker's theory showed a way of reducing TPZSGs to *n*-person zero-sum games.

N-person Non-zero-sum Games

But even this idea of a coalition-borne solution is unavailable in an *n*-person non-zero-sum non-cooperative game, which in normal form depicts a decision-making process not much different from that captured by 'bimatrix' games,[27] with the difference that now interacting decision-makers (players) are n (> 2). Apart from that, decisions are again taken autonomously, without negotiation, and out of a limited set of alternatives for each player. There being more than two players operating, 'a matrix formulation on the plane is not possible for such games', which renders the chance of a portrayal of possible outcomes and visualization of equilibrium strategies quite remote. A given *n*-person game may admit more than one admissible Nash equilibrium solution. The prefix non-cooperative only signifies those rules which 'forbid the negotiation of binding and enforceable agreements', rather than its pay-off structure. In the absence of negotiation, coalition formation is naturally impossible. Although non-cooperative solutions to *n*-person games are exposed to 'all of the objections raised against two-person mixed-motive games', they can prove highly 'illuminating'.[28]

APPLICATIONS AND APPLICABILITY OF GAME THEORY TO IR

Applicability and Limitations of Game Models: PD and Chicken

PD and chicken are two models of game theory hailed as generic metaphors for international politics and holders of great potential in the analysis of IR. PD is a 2 × 2 game, already classified under TPNZSG, coming with various explicatory tales and matrices but having the same basic substance. In the version of Albert Tucker, thesis adviser of Nash, two gangsters, Joe and Fox, arrested for a serious offence and given two hours to confess to or deny charges levelled by the district attorney, face the dilemma that if both confess, each will get 10-year prison sentence; if Joe confesses but Fox denies, Joe will be rewarded with just 2 year sentence and Fox will get 20 years, but successfully denying the charges both can get away with 4 years. Each prisoner is placed in separate rooms with no chance to communicate with each other. Their assessments of the four possible outcomes are given in Figure 11.3, with numbers denoting prison terms. The insightful point is that both Joe and Fox would be rationally inclined to confess to their crimes, no matter what option the other one settles for, the reason being: if Joe also confesses, then 10 years in prison for Fox is better than 20 and should Joe deny the charge, then Fox's 1 year is surely better than 2. Going through an analogous way, Fox would also find confessing a better option, regardless of what Joe does. This is somewhat tragic as well as 'counter-intuitive', since both know that if both could deny the charges, both would get just a mild 2-year sentence, incomparably better than 10. But being fully rational, none seem to be able to afford this 'intuitively logical conclusion'.[29]

What is to be noted is that the difficulties of the dilemma do not much depend on the distance between the rewards and punishments. Suppose public prosecutor makes the jail terms none, 1 day, 40 years and 45 years or makes it as absurd as follows.

	Joe Does* Not *Confess	***Joe* Does *Confess***
Fox *does not* confess	Acquittal + Euro 1,000	Slow Death through Third Degree
Fox *does* confess	Acquittal + Euro 10,000	Fast But Painful Death

Source: Adapted from Campbell, 'Introduction: Background for the Uninitiated', p. 6.

		Joe: Deny	Joe: Confess
Fox	Deny	4, 4	–20, –2
	Confess	–2, –20	–10, –10

FIGURE 11.3 Prisoner's Dilemma

Source: Adapted from Martin Peterson, *The Prisoner's Dilemma* (Cambridge: Cambridge University Press 2015), 2.

The same logic would compel both Joe and Fox to think of confessing, letting go acquittal with a tidy $1,000 in purse for quick but painful death, just because both are rational.[30] Or consider Figure 11.4, in which there are two strategies, termed C (cooperate) and D (defect), for each player. The players' preferences about the outcomes are expressed in the boxes, higher numbers betokening more preferred outcomes.[31] If both cooperate, both of them get 3, 3 or two years. But tempted to get 1, 4 or 4, 1 both defect, with each of them getting 2, 2. Such is the fate of rationality.

Whether in the most famous version from James Dean's classic 1955 film *Rebel without a Cause* or in others, chicken, another 2×2 game, consists two teenagers driving two automobiles to a cliff's edge, or towards each other, without swerving, although logically one has to swerve or both may die in the crash. Anyway, the swerver will be called a chicken. To give themselves handicaps that their resolve not to swerve credible, one player may wear sunglasses, while the other may throw out beer bottles from the car or may even tear out the steering wheel from it. With numerical pay-offs, the game will look like as shown in Figure 11.5.

If we put the game in cooperate/defect format like PD, the pay-off matrix will look like as in Figure 11.6. If both players cooperate (read swerve), both get 0, 0. If anyone defects, means goes straight, the pay-offs are as in the top right-hand and bottom left-hand boxes. If both defect, means go straight and crash, the results are shown in the bottom right-hand box.

An analogy between PD (despite its counter-intuitiveness) and practical international politics has been provided by many, including Campbell. Even after entering a

		Player 2 C	Player 2 D
Player 1	C	(4, 4)	(1, 5)
	D	(5, 1)	(2, 2)

FIGURE 11.4 Prisoner's Dilemma in CD Format

Source: Adapted from James D. Morrow, 'The Strategic Choice of Signalling, Commitment, and Negotiation in International Politics', in *Strategic Choice and International Relations*, eds David A. Lake and Robert Powell (Princeton, NJ: Princeton University Press, 1999), 81.

	Swerve	Straight
Swerve	0, 0	–2, +2
Straight	+2, –2	–10, –10

FIGURE 11.5 Chicken with Numerical Pay-off Matrix

Source: Adapted from James D. Morrow, 'The Strategic Choice of Signalling, Commitment, and Negotiation in International Politics', in *Strategic Choice and International Relations*, eds David A. Lake and Robert Powell (Princeton, NJ: Princeton University Press, 1999), 81.

		Player 2 C	Player 2 D
Player 1	C	0, 0	–2, +2
	D	+2, –2	–10, –10

FIGURE 11.6 Chicken in CD Format

Source: Adapted from James D. Morrow, 'The Strategic Choice of Signalling, Commitment, and Negotiation in International Politics', in *Strategic Choice and International Relations*, eds David A. Lake and Robert Powell (Princeton, NJ: Princeton University Press, 1999), 81.

pact on nuclear armament to prevent a mutually destructive holocaust, the USA and the USSR face the problem that most elaborate inspection safeguards built into the pact may not be equal to preventing one of the sides from secretly rearming, to the extent concealable from the other, and extract impermissible concessions and advantages. So since each side considers its own nuclear superiority as the ideal state, and its vulnerability as far worse than mutual destruction, each would mentally argue that its own possession of unmixed nuclear superiority is a better safeguard of world peace than mutual nuclear disarmament and would justify its breach of the pact as a purely defensive reaction. The dilemma is whether each would betray the pact to rearm secretly and the uncertainty situation is as follows.

	Player 2 Sticks to the Pact	*Player 2 Betrays the Pact*
Player 1 Sticks	No mutual destruction	Worse fate than mutual destruction
Player 1 Betrays	The ideal outcome	Mutual Destruction

Source: Richmond Campbell, 'Introduction: Background for the Uninitiated', in *Paradoxes of Rationality: Prisoner's Dilemma and Newcomb's Problem*, eds Richmond Campbell and Lanning Sowden (Vancouver: University of British Columbia Press, 1985), 7.

Campbell finds this version of PD quite in sync with reality, where true to its logic in previous situations, each of the superpowers should betray the agreement. An alternative to the PD type arms race model may be one where each side sincerely prefers mutual disarmament over anything else but believes that the other side is subject to PD logic. So while really wishing to disarm, it believes the other is bound to cheat. So cooperation is curbed by mutual distrust.[32]

PD has been used in other international contexts as well. For example, Steven Brams has applied it as a bargaining model in the Yom Kippur War (1973) to show why the superpowers did not come to fisticuffs during its course, but by employing a concept of equilibrium different from Nash.[33]

Chicken has been applied mainly to understand the dynamics of the 1962 Cuban Missile Crisis, known to all students of IR. A standard application is shown in Figure 11.7.

Here, the two players, 'White House' and 'Kremlin', have only two choices—'back down' or 'stand firm'. Their preference rankings (ones for the 'White House' kept on

		Kremlin	
		Wimp Out	Hold Out
White House	Wimp Out	3, 3	2, 4
	Hold Out	4, 2	1, 1

FIGURE 11.7 Cuban Missile Crisis in Chicken

Source: Adapted from Peter G. Bennett, 'Modelling Decisions in International Relations: Game Theory and Beyond', *Mershon International Studies Review* 39, no. 1 (April 1995): 24.

the left in each pair) range from 4 (best) to 1 (worst). Even though the two sides clearly do not share the same preferences, the game does not represent a clear and simple zero-sum situation. If neither side backs down, then both would meet the worst possible outcome, that is, certain death (1, 1).

Conclusions drawn from this game arising from irresponsible teenage desire for honour have been considered relevant to any situation—irrespective of context—where there are two sides with similar options. But there is a knotty problem of choice here. If one assumes that the other side will choose in the same way as oneself, then both the choices are destined to prove 'wrong'. If the other side swerves, then one ought to drive on, and vice versa. Analysis of the game reveals two (Nash) equilibria: '4, 2' and '2, 4'. Thus, a 'win' for either side seems stable because the loser could only move to an even worse outcome. Thomas Schelling (whom we have already dealt with in Chapter 6) has given an elaborate account of the tactics to be expected, many of which are as observable in our quotidian experiences as in high politics.[34]

Contextualizing the application of chicken in the Cuban crisis, the US policymakers, aiming at quick removal of the Soviet missiles about to be deployed in Cuba, contemplated two alternative courses of action: (a) a naval blockade (B), euphemistically called 'quarantine', to forestall the transportation of more missiles, coupled with sterner measures to induce the Soviet Union to withdraw missiles already planted; (b) a surgical air strike (A) to decimate the missiles already positioned, supplemented by a possible invasion of the island. The policy options the Soviet Union pondered were: (a) withdrawal and (b) maintenance of the USSR missiles. The array of choices, with 4 as the best and 1 as the worst, would be as shown in Figure 11.8.

Of course, this is a very skeletal, barebones portrayal of the 13-day crisis and does not mention many other options and choices considered by the superpowers. Besides, there were no ways to test that the outcomes mentioned in Figure 11.8 were the most probable prized ones or prized in terms of the game of chicken. For instance, if the Soviet Union really regarded an air strike on their missiles as detrimental to their most crucial national interests then the *AW* outcome could easily culminate into a nuclear duel between the sides, making it indistinguishable in value from *AM*. Another unwarranted simplification was making the game 'normal form', where the sides select their strategies simultaneously, although in actuality 'a continual exchange of messages, occasionally backed up by actions, occurred in those fateful days of October'.

		Moscow	
		Withdrawal (W)	Maintenance (M)
Washington	Blockade (B)	(3, 3) Compromise	(2, 4) Soviet victory US defeat
	Air strike (A)	(4, 2) US victory Soviet defeat	(1, 1) Nuclear war

FIGURE 11.8 Cuban Missile Crisis in Chicken

Source: Adapted from Steven J. Brams, *Negotiation Games: Applying Game Theory to Bargaining and Arbitration* (London: Routledge, 1990), 105.

.8	.8	.8	.8	.8	.8	.8	.8	.8
.8	.64	.512	.409	.328	.262	.210	.168	.134

FIGURE 11.9 Decreasing Credibility of Threat Strategies in Chicken

Source: Adapted from Lieber, Theory and World Politics, p. 27.

Quite apart from these problems, even though both sides were deemed to be on the collision course in the popular perception, most of the informed observers concur that both sides were wary of taking any hasty, 'irreversible' steps as chicken drivers are sometimes prone to do. And Brams as well others have shown that if the chicken game is played in 'iterated' contexts, not in normal form but in extensive form in sequential contexts, and more as a 'theory of moves' than as a theory of games, then the strategies as well as the matrices may undergo significant changes. These modifications are particularly important in the context of dynamic games, which characterize IR.[35] There the players are not irresponsible teenagers, who go home to their parents' care after racing cars, oblivious of the fact that the credibility of even the most successful player's threat strategies at each new game may decrease in value at every later move in the order as shown in Figure 11.9.

These words apply equally to PD. Had PD been played repeatedly, astute players could establish a pattern of previous choices by communicating each other in a way that would reward the adoption of cooperative strategy. But where the game ends after *n* rounds and the final outcome is accordingly definitely known, cooperation on this round is not profitable since, 'with no plays to follow, the players are in effect in the same position if they played the game only once'.[36] But where the game is played for many rounds, a player possessing threat power could decide to play cooperatively in the first round and then switch to retaliation in probable future rounds if the second player did not care to reciprocate. So long as both acted on the assumption that the game is to be repeated, the threat remains effective and both players cooperate. So a host of cooperative equilibria can be established in repeated games that are inconceivable in the single-shot editions of the same games.[37]

It is true that just like chicken, PD has also illumined many other contexts of IR. Duncan Snidal says that analysis through the lens of PD has thrown incisive insights into important negotiations as General Agreement on Tariffs and Trade (GATT) or Strategic Arms Limitation Talks (SALT), which the unaided capability of historical-archival research could never match. And extending its logic and that of collective action to IR reveals why international cooperation is not sometimes forthcoming even when it serves the interest of all states.[38] But Snidal himself contends that to understand the applicability of game theory to IR as a *theory* we have to move beyond fitting 'realist' game models such as chicken and PD to discrete episodes of international politics, diagnose their problems and find ways of overcoming them. In the next section, we will see how this can be done.

Applicability of Game Theory in IR beyond Game Models

Some scholars plead for the general applicability of game theory to IR through broadening of the domain of rational actor models, beyond the constraining limits of the traditional realist paradigm in order to grasp a more complex world preoccupied as much with issues of conflict as with those of cooperation. Snidal suggests 'use of game models to understand different aspects of international politics in terms of a unified theory', not its underuse as a 'descriptive rather than analytical tool…in the spirit of sorting out whether the Cuban Missile Crisis was really chicken or PD'. Rather than lying in 'reconstructing and interpreting particular events', the primary benefit of game theory supposedly consists of 'redescribing' our world, as done in Snyder and Diesing's application of game models to 16 historical cases, which exhume the game structure of each crisis through an elaborate historical exposition of each, athwart sharply different accounts of these crises.[39]

However, a holistic application of game theory has to move these endeavours which show more virtuosity 'in reconstructing a crisis in game terminology' than demonstrating the capacity of the theory properly. Since the real potential of game theory in both empirical and theoretical areas is best revealed when it is employed to create new insights rather than when illumining past episodes, its subject should be 'the goal-seeking behaviour of states in an interdependent international system', not attempts to forecast outcomes resulting from their 'non-purposive or non-systematic behaviour'. For, if the underpinning assumption of international politics is 'self-interested action by strategically rational states', whose preferences and strategies can be divined and pay-offs calculated accordingly, then game theory can produce important verifiable prophesies and propositions. Snidal shows how the possibilities of applying game theory to international politics as *theory* should distinguish the diverse logics of its applications as metaphor, analogy and model and their limitations.

Although *metaphors* are helpful in facilitating movement of ideas across discrete domains and, thus, acquire great heuristic and expository value, just because they are also susceptible to misuse or clumsy misapplication, 'metaphorical richness must be progressively restricted by more precise formulations as research advances'. For example, the security dilemma in international politics captured by the metaphor

Hobbesian state of nature should be restricted from leading to the conclusion that cooperation away from present international anarchy is possible only with a Hobbesian international sovereign. Snidal reminds proponents of metaphor-based application of game theory that just the offhand branding of an issue by a game name does not facilitate greater use of its explanatory power, but only provides an embellished restatement of known facts.

Snidal reminds analogists that their inferences are 'tentative until empirically confirmed'. For example, the current fashion of tying international politics to neoclassical microeconomic models revolves around an equation of states in the international system and firms in an oligopoly, and implicit belief in the game-theoretic structure underlying that market. The logic is as follows.

Economic marketplace ←→ International system
Firm ←→ Nation State
Firms maximize profits ←→ States maximize survival
Oligopolists ←→ Great powers
Market concentration ←→ Concentration of powers
Price wars ←→ Military wars
Both are self-help systems
Both firms and states act strategically

Since this *analogy* transfers inferences from oligopoly theory to the international system to contend that just as oligopolistic market concentration generates market stability and decreases price wars, so also concentration of power in the international system make for greater system stability and lesser conflicts, it is contingent for its value as a hypothesis on how compelling the initial correspondences are. These would be thin because of these limitations of analogies: (a) even identified disanalogies cannot be stamped out, (b) interesting aspects not covered by the analogy cannot be scrutinized and (c) the inferential logic is primarily inductive, with little room for deductive logic, which is the forte of the game-theoretic approach. So models are better substitutes for directly incorporating the most salient features of the international system.

A plea for *models* as a better option for application of game theory to international politics rests on the crucial distinction between 'models of processes or of things' (e.g., of 'a particular arms race') and 'models of theories' (like arms race as a generic category of phenomena). The former entails 'abstraction of an entity's properties' to portray them in a simpler set of relations. But the model of a theory comprises a bunch of 'linked law-like statements' relevant for a range of phenomena. Game theory highlights 'a healthy tension' between these two sorts of models.

Coming to *theory*, since the same representation is able to serve as a model for different theories, exposition of a model will rest on the deductive structure of the theory it is rooted in, as also on interpretation of its fundamental assumptions and theoretical constructs. As a *theory* of calculated and deliberate intentional behaviour, game theory's presupposition of rationality permits analysis of solutions in light of the premeditated behaviour of actors. Particularization of different political-institutional environments, like capitalist market in economic oligopoly theory pitted against anarchic international society in BOP theory, fixes rules of the game that lead to 'different interpretations of models' and finally to different models, after the rules are brought more openly into the analysis. Or else, different stipulations of actors' policy choices and/or preferences may ensure that different games (e.g., chicken versus PD) become salient within the same theoretical interpretation of the international system. But away from its enormous diversity of models, to become a general theory of international politics instead of a general theory of strategic behaviour, game theory needs its own specific empirical assumptions. For instance, by accepting power-maximizing states as the principal actors, game theory embraces the realist position, although its basic approach does not accord well with realism. Its presumption that actors are rational never necessitates that the key actors are states or that they seek power above everything else. With a different take of these assumptions, game theory is quite comfortable with a modified structural approach. Bringing battling perspectives into one single framework and delineating their different empirical assumptions, game theory does not have to remain confined to the realist paradigm, even though it assumes 'goal-seeking behaviour in the absence of centralized, authoritative institutions', and, thus, throws light on the fundamental issues of international anarchy and the implications of different configurations of national interests and political circumstances for international conflict and cooperation. Its simplifying assumptions broaden our understanding sufficiently to grasp the profounder interpretations of international politics.

Snidal thinks that game theoretic research has gained sufficient maturity to use metaphors and analogies sparingly and go for the rigorousness and deductive power of the model to benefit from interpretative depth and space of the theoretical framework in which it is embedded in. After tightening up the correspondences between empirical situations and separating assumptions from predictions, a transformed game theory would not just remain 'a new language in which to rewrite history or to restate our arguments' but would prove 'a powerful tool for expanding our understanding and for stimulating research'. Its only price is efforts to clarify the assumptions and link the deductive logic to empirical reality.

The plea for a game theoretic transformation of international politics draws additional strength from the fact that its basic concepts—such as *strategies*, *strategic rationality*, *preferences* and *pay-offs*—supply a rational map for constructing theory in IR not just for simple 2 × 2 games but other game models as well. We can consider strategies 'as simplified representations of general policy stances'. For instance, in trade negotiations it makes sense to compare strategies of 'free' international trade against 'restricted' international trade, without bothering about the nitty-gritty of dissimilar treatment of steel and textiles, or the issues of tariff and non-tariff barriers. And strategy talks of curtailing military spending, of lessening international tensions or of furthering environmental protection make eminent sense regardless of if someone has made an elaborate list of the ways to implement these policies under every conceivable contingency. The simple 2 × 2 game pushes this quest to its logical extreme by providing only two choices, often simplistically labelled 'cooperate' and 'do not cooperate'. In spite of limitations, this useful simplification reveals the dynamics and dilemmas of an issue area. And game theory is useful because some of its most productive findings have come 'when dynamic problems are treated as static choices of strategies which actors will play through time'.

Coming to strategic rationality, at the heart of a game-theoretic interpretation, some of the best critiques of the realist assumption of states as rational actors such as 'bureaucratic politics, psychological models of decision-making, social choice and complex organizations' offer lessons that could be best tackled by game-theoretic analysis. But game theory also interrogates the realist conception of non-strategic rationality, which, revolving around the anarchy problematique, envisions states fending for themselves as they pursue their contradictory interests. Through back calculation from the conflictive nature of this 'self-help' environment, the situation is misperceived as a zero-sum one that leaves no scope for cooperation. For game theory, this scenario is valid only for artificially constructed, two-player parlour games. For, in real international politics, interests of states cannot be pursued amidst such pure opposition of interests. Against these, a strategic rationality embodies the realization that even furtherance of egoistic interest needs judging one state's choices against choices of other states. The corollary of this is that national policymakers have to avail of opportunities for cooperative interactions even when seeking protection amidst conflictual interactions.

Snidal argues that game-theoretic analysis facilitates consideration of two important aspects of rationality. The first, found in both non-strategic and strategic conceptions of rationality, relates to the ability to eschew short-term benefits for longer-term ones. The second, more definitive of strategic rationality, holds that person's choice of courses of action is dictated by their preferences and expectations of how others will behave. Thus, when a state embarks on a certain action, it does not necessarily imply that its immediate outcome is a preferred state of affairs for that state. It could have been a strategy gambit for some other ulterior objective. The game-theoretic presupposition of strategic rationality foregrounds the rational choice of state policy, thereby permitting autonomy in state choice even when predicting and explaining those choices deterministically in terms of the overall strategic interaction. In this way, the game model links purposeful behaviour with a glimpse of the structure of international politics which constrains that behaviour. It links systematic macro-theory with voluntaristic decisions. Even when states possess choices, these are dictated a greater or lesser extent by the contingencies of international politics.[40]

CRITIQUE AND EVALUATION OF GAME THEORY

But there are critics who do not see any future for a thorough overhauling of IR theory through a theory of games because of its supposedly crippling weaknesses. Forward described five main strands of criticism of game theory in IR literature, namely 'indeterminate, impracticable, static, preposterous and irrelevant'. Of them, indeterminacy targets mainly 'the incoherence and ambiguity of the solutions of non-zero-sum games, which must surely be the model of political situations if any games are'. Impracticability concerns the assertion of many critics including Herbert Simon that 'the fantastic information processing and reasoning powers expected of the players by the assumptions about knowledge and rationality' remove the theory from any truck with actual decision-making. It is termed static for the widely held belief that game theory's practical utility would be minimal 'unless it enters the time dimension in one way or another'. As a game is represented in present theory, it is 'once for all'. Values in the matrix are from the start fixed quantities and given. Many theorists accordingly 'favour the concept of a supergame or sequence of games in the course of which a player's evaluation of different outcomes evolve, so do his/her estimates of the other players' values....' Preposterousness emphasizes the sense that 'the inputs (i.e., the evaluation of all the possible outcomes by a player) are not available in real life until after the decisions are made'. Since the 'decisions in fact precede the ordering', the only way to understand the ranking of utilities by the players is to observe the decisions. So game theory is branded as 'a topsy-turvy model, since it represents things as happening in the opposite order from what they really do'. Forward offers a less paradoxical version of the same point of view that 'in social and political situations it is impossible to assess the costs and pay-offs'. The charge of irrelevance is borrowed from P. M. S. Blackett, who said that 'the influence of the Theory of Games has been almost wholly detrimental...[It is] a branch of pure mathematics and almost wholly irrelevant to decisionmaking'.[41]

Because of the charge of preposterousness, the only way Forward sees of fitting a negotiation into the framework of game theory is by regarding 'all the real bargaining as taking place before the game starts and as being concerned largely with determining what game is to be played. The play of the game itself is tautological and its analysis of no interest'. He puts Anatol Rapoport's emphatic rejection of game theory as an aid to decision-making into the 'irrelevant' box, 'though it is specifically the attempt to use game theory prescriptively that offends him'.[42]

Some other criticisms of game theory revolve around how it oversimplifies reality, contains logical errors and has an understanding of rationality which is unsuitable in IR. Regarding the oversimplification of game-theoretic models, the most frequent objection is that game theorists 'squeeze' the world into a largely static matrix, which does not accommodate additional information that may be needed to understand the situation, 'such as details about the context of interaction, insights into the personalities and behaviour of decision-makers, understandings about the diplomatic or foreign policy process, issues that may be linked to the issue in question and differing perceptions of the players'. Introducing these dimensions and facets and turning 2×2 games into n-person games with $n > 2$ will require intimidating mathematical knowledge for most social scientists.[43]

Brams points to two other drawbacks that inhere in the logic of conventional game theory in IR, namely 'mis-specifying the rules' and 'confusing goals with rational choice'. Since a game is 'the totality of rules that describe it' and prescribe behaviour, 'any game-theoretic model should propose rules of play that reflect how players think and act in the strategic situation being modelled'. In terms of this reality check, standard game theory often 'misses the mark' because players do not 'choose strategies simultaneously as assumed in the normal or strategic form of a game...represented by a pay-off matrix' and neither do they 'adhere to a specified sequence of choices as assumed in the extensive form of a game...represented by a game tree'. Contrarily, often play resumes 'at some initial state or status quo point', where players strive to ascertain if they can improve their position by moving or staying put as a result of which normally accepted rational postulates are often flouted for both types of game. For instance, players sometimes employ 'dominated strategies' in PD 'or do not use backward induction', beginning from 'endpoints of a game tree', when PD is repeated or the number of rounds is fixed. Regarding the second drawback, Brams says that while rationality is rightly 'applied to the efficiency or efficacy of the means or instruments used for desired ends', it 'does *not* concern the ends themselves, which are neither rational nor irrational'. While the way people 'come to harbour the goals that they do' is a legitimate development question, that is not germane to the idea of rationality game theory espouses'.[44]

Mesquita lists three weaknesses of game theory that he says are resolvable with time, two of which I have touched upon. First, it is 'demanding in terms of information assumptions'. This does not mean that 'every bit of information in a game-theoretic setting must be known by everyone', since game theory has always accommodated problems of uncertainty 'in the guise of incomplete and/or imperfect information'. This requirement rather signifies that 'games that involve learning must start with the premise that everyone knows everyone else's prior (initial) beliefs (though not their prior knowledge which can be private)'. Second, the multiple equilibria that even many simple games like chicken or BOTS possess imply that game theory can 'predict a probability distribution over actions but cannot say definitively when one action or another will be taken provided such action is part of an equilibrium strategy'. Third, although all players in game theory 'are assumed to play with equal skill', each does not play the game in the same way, making the concept of player types salient in game-theoretic situations of limited information, particularly where there is no clear idea of what outcomes other value or how much. Signalling games have emerged as a sound way to address these problems. And one front-ranking approach that addresses the uneven distribution of skills is the concept and models of bounded rationality. But they do not fully address these problems, because they 'proliferate the range of actions predicted to be possible'.[45]

From post-structuralist and communicative action perspectives, game theory has been criticized for perpetuating the construction of realist discourse. For Hurwitz, depicting wars and conflicts as strategic games that are to be won and lost 'forecloses perceiving of cooperation as an action possibility'.[46]

Sen has exposed the faulty logic of PD, 'often treated, with some justice, as the classic case of failure of individualistic rationality', whose two strategies he terms as 'selfish and unselfish' (in our language confess and deny). In his rendering, the presumption of PD is that:

> Each player is better off personally by playing the selfish strategy no matter what the other does, but both are better off if both choose the unselfish rather than the selfish strategy. It is individually optimal to do the selfish thing: one can only affect one's own action and not that of the other, and given the other's strategy—no matter what—each player is better off being selfish. But this combination of selfish strategies…produces an outcome that is worse for both than the result of both choosing the unselfish strategy. It can be shown that this conflict can exist even if the game is repeated many times.

Sen argues that not only in real life but 'even in controlled experiments in laboratory conditions', people playing PD frequently pursue the unselfish course. While interpreting these anomalous experimental results, the game theorist is prone to attribute it to the poor intelligence of the players. But for Sen, a more productive approach might consist in allowing 'the possibility that the person is more sophisticated than the theory allows and that he has asked himself what type of preference he would like the other player to have, and on somewhat Kantian grounds has considered the case for himself having those preferences, or behaving as if he had them'.[47]

Against these criticisms, enthusiasts mention the achievements of game theory in areas of 'alliance formation, reliability and termination'; deterrence theory; the development of nuclear and conventional strategy throughout the Cold War; the current research on economic sanctions as a tool of diplomacy and limitations of sanctions as strategy; the growing literature on paths to cooperation in IR 'directly out of repeated games'; enlightenment about factors facilitating 'conflict initiation, escalation and termination' and insights provided in many other areas.[48]

CONCLUSION

This chapter has discussed the contextual origins of game theory in lack of direct communication between the two superpowers during the Cold War, and disciplinarily in that both IR and game theories try to grasp the problems of conflict and cooperation. After showing how in the theory TPZSGs are resolvable through the minimax theorem, TPNZSGs are tractable through Nash equilibria and mixed strategies up to a point before the players establish their negotiation sets where the cooperative stage is no more amenable to mathematics, *n*-person games are made applicable to politics through the concepts of Shapley value and theory of coalitions, I discussed how TPNZSGs of PD and chicken have been used to understand international processes like the USA–USSR arms race and their decision-making during the Cuban Missile Crisis. After this followed scholarly suggestions about how game theory can be applied to international politics beyond metaphors, analogies and its models, and as theory. Finally, I embarked upon the limitations of game theory and its achievements. I balance the last two with the help of this insight from Heap and Vardoulakis, with which I end: 'understanding why game theory does not, in the end, constitute the science of society (even though it comes close) is terribly important in understanding the nature and complexity of social processes',[49] among which I place international relations.

REVIEW QUESTIONS

1. Describe the contextual and disciplinary origins of game theory.
2. Identify the most basic terms of game theory and give their meanings.
3. Enumerate the basic types of games and show how TPZSGs are mathematically tractable.
4. Explain the basic characteristics of TPNZSGs with special reference to PD and BOTS.
5. Analyse how *n*-person zero-sum games with $n > 2$ are made mathematically tractable through the concept of Shapley value and Riker's theory of coalitions.
6. 'A well-known example of the two-person non-zero-sum non-cooperative game is PD.' Examine the statement.
7. Examine the applicability of chicken as a model of IR.
8. Attempt an evaluation of game theory.

NOTES AND REFERENCES

1. Lieber, *Theory and World Politics*, 18–19.
2. Pierre Allan and Cédric Dupont, 'International Relations Theory and Game Theory: Baroque Modelling Choices and Empirical Robustness', *International Political Science Review* 20, no. 1 (January 1999): 23–47, here 23.
3. James D. Morrow, 'Game Theory', in *Encyclopedia of International Relations and Global Politics*, ed. Martin Griffiths (London: Routledge, 2005), 296.
4. Andrew H. Kydd, *International Relations Theory: The Game-theoretic Approach* (Cambridge: Cambridge University Press, 2015), 6–7.
5. Theodore L. Turocy and Bernhard von Stengel, 'Game Theory', CDAM Research Report LSE-CDAM-2001-09 (8 October 2001): 4.
6. Roger B. Myerson, *Game Theory: Analysis of Conflict* (Cambridge, MA: Harvard University Press, 2007), 1–4.
7. Nigel Forward, *The Field of Nations: An Account of Some New Approaches to International Relations* (New York, NY: Palgrave Macmillan, 1971/2016), 24–26.
8. Peter G. Bennett, 'Modelling Decisions in International Relations: Game Theory and Beyond', *Mershon International Studies Review* 39, no. 1 (April 1995): 19–52, 20–21.
9. Takako Fujiwara-Greve, *Non-cooperative Game Theory* (Tokyo: Springer, 2015), 2. Also see Turocy and Stengel, *Game Theory*, 6; Forward, *Field of Nations*, 38–39.
10. Randall W. Stone, 'The Use or Abuse of Game Theory in International Relations: The Theory of Moves', *The Journal of Conflict Resolution* 45, no. 2 (April 2001): 216–244, esp. p. 217; Turocy and Stengel, *Game Theory*, 30–31.
11. Neumann cited in Turocy and Stengel, *Game Theory*, 3.
12. Neumann cited in E. van Damme, 'Refinements of the Nash Equilibrium Concept', in *Lecture Notes in Economics and Mathematical Sciences, 219*, eds M. Beckmann and W. Krelle (Berlin: Springer-Verlagg, 1983), 13.
13. Beautiful in the sense in which physicist Paul Dirac described Einstein's work on gravitation as 'search for a beautiful theory'. See Roberto Franzosi, *From Words to Numbers: Narrative, Data, and Data and Social Science* (Cambridge: Cambridge University Press 2004), 126.
14. Forward, *Field of Nations*, 27–29.
15. Frank C. Zagare, *Game Theory: Concepts and Applications* (Beverly Hills, CA: SAGE Publications, 1984), 33–34.
16. Forward, *Field of Nations*, 30.
17. R. Duncan Luce and Howard Raiffa, *Games and Decisions: Introduction and Critical Survey* (New York, NY: Dover Publications, 1957/1985), 88–92.
18. Forward, *Field of Nations*, 36–39.
19. Myerson, *Game Theory*, 98, 108–109.
20. Anatol Rapoport, *N-Person Game Theory: Concepts and Applications* (Mineola, NY: Dover Publications, 1970), 158.
21. L. C. Thomas, *Games, Theory and Applications* (Mineola, NY: Dover Publications, 1986), 86.
22. Andrew M. Colman, *Game Theory and Its Applications in the Social and Biological Sciences*, 2nd ed. (London: Routledge, 1995), 162.
23. Forward, *Field of Nations*, 33–34.
24. Luce and Raiffa, *Games and Decisions*, 156.
25. Colman, *Game Theory*, 172; Forward, *Field of Nations*, 34.
26. Kaare Strøm, 'Coalition Theory and Government Formation', in *The Oxford Handbook of Comparative Politics*, eds Carles Boix and Susan Carol Stokes (Oxford: Oxford University Press, 2007), 789–792.

27. Bimatrix games are essentially two-person non-zero-sum games.
28. Tamer Basar and Geert Jan Olsder, *Dynamic Noncooperative Game* (New York, NY: SIAM, 1999), 88–90; Colman, *Game Theory*, 162.
29. Martin Peterson, 'Introduction', in *Prisoner's Dilemma* (Cambridge: Cambridge University Press 2015), 1–2.
30. Richmond Campbell, 'Introduction: Background for the Uninitiated', in *Paradoxes of Rationality: Prisoner's Dilemma and Newcomb's Problem*, eds Richmond Campbell and Lanning Sowden (Vancouver: University of British Columbia Press, 1985), 4–7.
31. James D. Morrow, 'The Strategic Choice of Signalling, Commitment, and Negotiation in International Politics', in *Strategic Choice and International Relations*, eds David A. Lake and Robert Powell (Princeton, NJ: Princeton University Press, 1999), 80–81.
32. Campbell, 'Introduction', 7; Bennett, 'Modelling Decisions in International Relations', 31.
33. Steven J. Brams, *Negotiation Games: Applying Game Theory to Bargaining and Arbitration* (London: Routledge, 1990), 104.
34. Bennett, 'Modelling Decisions in International Relations', 23–24.
35. Brams, *Negotiation Games*, 105–116; Randal Stone, 'The Use and Abuse of Game Theory in International Relations: The Theory of Moves', *The Journal of Conflict Resolution* 45, no. 2 (April 2001): 216–233.
36. Brams, *Negotiation Games*, 120.
37. Steven J. Brams, *Theory of Moves* (Cambridge: Cambridge University Press, 1996), 129, 155.
38. Duncan Snidal, 'The Game Theory of International Politics', *World Politics* 38, no. 2 (October 1985): 25–57, here 28. The article is also available in Kenneth A. Oye, ed., *Cooperation under Anarchy* (Princeton, NJ: Princeton University Press, 1985), 25–57.
39. Glenn Snyder and Paul Diesing, *Conflict among Nations* (Princeton, NJ: Princeton University Press, 1977).
40. Snidal, 'Game Theory', 29–40.
41. For the points and the cited sources see Forward, *Field of Nations*, 43–49.
42. Anatol Rapoport, *Strategy and Defense* (Manhattan, NY: Harper and Row, 1964), esp. Chapter 8, cited in ibid., 50.
43. Thomas Diez, Ingvild Bode, and Aleksandra Fernandes da Costa, eds, 'Game Theory', in *Key Concepts in International Relations* (London: SAGE Publications 2011), 72.
44. Steven J. Brams, 'Game Theory in Practice: Problems and Prospects in Applying It to International Relations', in *Millennial Reflections on International Studies*, eds Michael Brecher and Frank P. Harvey (Ann Arbor, MI: University of Michigan Press, 2002), 392–393.
45. Bruce Bueno de Mesquita, 'Accomplishments and Limitations of Game-theoretic Approach to International Relations', in *Millennial Reflections on International Studies*, eds Michael Brecher and Frank P. Harvey (Ann Arbor, MI: University of Michigan Press, 2002), 382–389.
46. Roger Hurwitz, 'Strategic and Social Fictions in the Prisoner's Dilemma', cited in Diez, Bode, and da Costa, 'Game Theory', 72.
47. Amartya Sen, 'Rational Fools: A Critique of the Behavioral Foundations of Economic Theory', *Philosophy and Public Affairs* 6, no. 4 (Summer, 1977): 317–344, here 340–341.
48. Mesquita, 'Accomplishments', 382.
49. Shaun P. Hargreaves Heap and Yanis Varoufakis, *Game Theory: A Critical Text* (London: Routledge, 2004), 2.