G

GAME THEORY

Game theory is the branch of decision theory that analyzes interdependent decision problems between rational, strategic agents. A rational agent is one who has a consistent set of preferences defined over some set of possible outcomes and who makes choices consistent with these preferences. A strategic agent is one who, given these preferences, reasons about the best course of action to take in order to satisfy them. Interdependent decision problems arise when the outcome for any particular agent depends upon the actions chosen by all of the agents; that is, when the optimal choice for an agent A depends upon the choices made by other agents, and the optimal choice for the other agents depends in turn upon the choice made by A. It is this strategic feature that distinguishes game-theoretic problems from simpler decision problems such as parametric choice under conditions of risk or uncertainty.

The birth of modern game theory is usually attributed to von Neumann and Morgenstern (1944). However, precursors to game-theoretic analyses of strategic problems can be found in Zermelo (1913), Borel ([1921] 1953), and von Neumann ([1928] 1959), as well as in the works of Hobbes and Hume.

A Theory of Utility

One of von Neumann and Morgenstern's primary contributions was their development of a mathematical theory of utility, allowing one to define, for a given agent, an interval utility measure unique up to a strictly increasing affine transformation. The need for such a notion of utility originates in the fact that in game theory, agents often need to make decisions under conditions of risk or uncertainty, and hence one needs a measure of how strong their preferences for a given outcome are.

If an agent's preferences over outcomes satisfy certain basic coherence criteria, it is possible to define a utility function with the property that if one makes choices consistent with one's preferences, one acts *as if* one were choosing to maximize expected utility. The following axioms (from Luce and Raiffa's [1957] classic text *Games and Decisions*) formalize the coherence criteria necessary to satisfy in order to define a von Neumann–Morgenstern utility function. Let $A = \{a_1, \ldots, a_n\}$ denote the set of outcomes, and let $a_j \precsim a_i$ denote that the agent either prefers a_i over a_j or is indifferent between them. A *lottery* $L = (p_1 a_1, \ldots, p_n a_n)$ is simply a randomization over outcomes, where the outcome a_i occurs with probability p_i. A *compound lottery* $Q = (q_1L_1, \ldots, q_mL_m)$ is a lottery over lotteries, where the chance that the lottery L_i occurs is q_i.

Ordering of Alternatives

For any outcomes a_i, a_j, and a_k either $a_i \precsim a_j$ or $a_j \precsim a_i$ (and possibly both). Moreover, the relation "$\precsim$" is *transitive*; that is, if $a_i \precsim a_j$ and $a_j \precsim a_k$, then $a_i \precsim a_k$.

Reduction of Compound Lotteries

Let $L^i = (p_1^i a_1, p_2^i a_2, \ldots, p_n^i a_n)$ be a lottery, for $i = 1, \ldots, m$. Then the agent is indifferent between the compound lottery $(q_1L^1, q_2L^2, \ldots, q_mL^m)$ and the simple lottery $(p_1a_1, \ldots, p_na_n)$, where $p_i = q_1p_i^1 + q_2p_i^2 + \ldots + q_mp_i^m$.

Continuity

Suppose that $a_n \precsim a_{n-1} \precsim \ldots \precsim a_1$. Then there exists a number u_i such that the agent is *indifferent* between a_i and the lottery $[u_ia_1, 0 \bullet a_2, \ldots, 0.\bullet a_{n-1}, (1 - u_i)a_n]$, which is denoted $\hat{a}_i$.

Substitutibility

In any lottery, $\hat{a}_i$ is substitutable for a_i.

Transitivity of Lotteries

The preference and indifference relations over lotteries are transitive relations.

Monotonicity

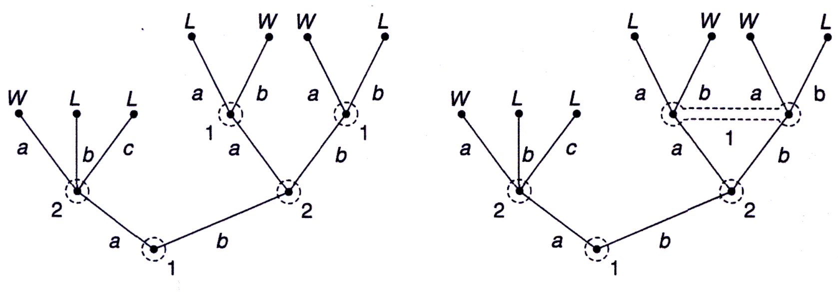
A lottery $(pa_1, (1 - p)a_n)$ is preferred or indifferent to $(p'a_1, (1 - p')a_n)$ if and only if $p \geq p'$.

If an agent's preferences satisfy the above axioms, it is possible to find a number u_i for each outcome a_i such that for any two lotteries L and L' the magnitudes of the expected values $p_1u_1 + \ldots + p_nu_n$ and $p'_1u_1 + \ldots + p'_nu_n$ indicate the preference between the lotteries. From this assignment of utilities to the basic alternatives, one can construct a utility function f over the set of risky alternatives (the lotteries). Consequently, when an agent makes choices consistent with her preferences, she acts as if she is choosing to maximize personal utility as measured by f.

Representations of a Game

Games are most commonly represented in an extensive or a strategic form. One also finds the strategic form referred to as the *normal* form, following von Neumann and Morgenstern, who believed that normally one should reduce the extensive form of a game to the strategic form for the purpose of analysis.

The extensive form uses a game tree to represent the order of play (see Figure 1). Each node in the tree represents a *choice point* for a particular player; the player whose turn it is to move at a particular choice point is indicated by a label attached to the node. All games have a privileged node, the *root* or *initial* node where the game begins. The leaves of the tree, also known as *terminal* nodes, represent endpoints, or outcomes, of the game. Every node in the game tree except for the terminal nodes has at least one edge lying on a path between it and a terminal node; such edges represent choices available to a player at that choice point. In some games, the moves available to a player depend not only on the previous moves of other players, but on the outcome of a chance event like the roll of a die. Such games may be represented by including a fictitious player in the game tree, Chance, whose available moves at a point correspond to the possible outcomes of the random event. A player's choice at a given point is a *move* in the game, and each edge has an attached label naming the move. A path from the root node to a terminal node is one possible *play* of the game. In Figure 1, terminal nodes are labeled with W or L, meaning that Player 1 wins or loses the game, respectively.



(a) Game of perfect information

(b) Game of imperfect information

Fig. 1. A simple two-player game in extensive form. Terminal nodes labeled "W" and "L" indicate whether Player 1 wins or loses, respectively.

If all players know their exact position in the game tree at every point, the game is said to be one of *perfect information*; all other games are of *imperfect* information. Although players do not always know their exact position in the game tree in a game of imperfect information, they often know that their position is one of a limited number of possible nodes. This subset of nodes is a player's information set. In an extensive form game, a player's *strategy* specifies the choices that the player would make at each of his information sets. A player's information sets are indicated in the game tree by grouping together those nodes among which the player cannot distinguish. Thus, an alternative definition of a game of perfect information is one in which all information sets contain only a single node. Figure 2b illustrates a game of imperfect information in which Player 1 moves first but keeps the move hidden from Player 2. When it is Player 2's turn to move, he does not know whether the choice occurs at the left or the right side of the game tree.

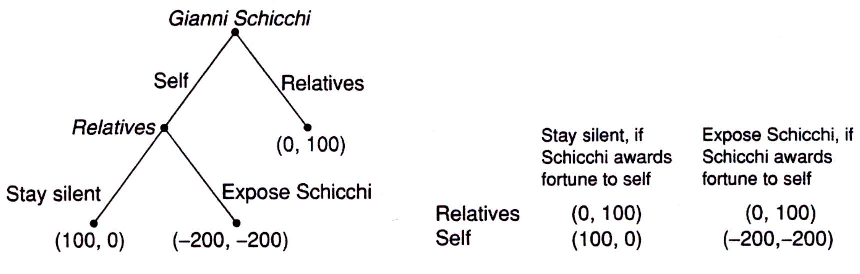
The strategic form of a game is a minimal representation that omits all information about the game except for the relationship between strategies and payoffs. The strategic form of a game consists of a set of players $P = \{1, \ldots, N\}$, a set of pure strategies S_i for each player $i \in P$, and, for each player i, a *payoff function* u_i that maps pure strategy profiles $\sigma = (\sigma_1, \ldots, \sigma_N) \in S_1 \times \ldots \times S_N$ to a real number r. In a two-player game, the strategic form can be represented as a matrix, where each row corresponds to a strategy for Player 1, each column a strategy for Player 2, and each cell the resulting payoffs obtained by Players 1 and 2 when they choose those respective strategies.

In many cases, it proves convenient to allow players to adopt *mixed* strategies, where they choose a pure strategy at random according to some probability distribution defined over the set of pure strategies S_i. The payoff for a mixed strategy $\bar{\sigma}_i$ is defined to be the expected payoff $\sum_\sigma P(\sigma \mid \bar{\sigma}_i) u_i(\sigma)$, where the sum is over all strategy profiles σ and $P(\sigma \mid \bar{\sigma}_i)$ denotes the probability that the strategy profile σ occurs when player i adopts the mixed strategy $\bar{\sigma}_i$.

Although it is often said that which form one uses to represent a game is merely a practical question, on the grounds that any game represented in one form may be represented in the other, this is a topic of some debate. To begin with, it is clear that moving from the extensive form to the strategic form results in a loss of information, for it is possible for two *different* extensive games to have the *same* strategic form. In some cases, this lost information may be relevant to the analysis of the game; if so, it may not always be possible to adequately analyze a game just given its strategic form (see Harper 1988). For example, Figure 2 illustrates the strategic and extensive forms for the decision problem central to Puccini's opera *Gianni Schicchi*, in which the causal dependencies between the players' choices are lost in the normal form, yet seem crucial to the game's analysis.

Noncooperative Games

In a noncooperative game, players independently decide what strategy to adopt in the light of their knowledge of the other players and the payoff matrix. Most of the classical results in game theory have been obtained for noncooperative games, for



(a) Gianni Schicchi, extensive form (b) Gianni Schicchi, strategic form

Fig. 2. Strategic and extensive form of the Gianni Schicchi. In Puccini's opera *Gianni Schicchi,* the wealthy Buoso Donati dies, and before his will is read, his relatives learn that he has willed a large portion of his fortune to friars. They conspire to have a noted mimic, Gianni Schicchi, impersonate Buoso Donati on his deathbed in order to dictate a new will. Gianni Schicchi agrees but while impersonating Buoso Donati and dictating a new will, he declares his wish to leave a large portion of his fortune to his devoted friend Gianni Schicchi. The relatives contemplate notifying the authorities but decide against it, knowing that the punishment for tampering with a will is banishment and amputation of a hand.

the inability of players to form coalitions and enter into binding agreements make noncooperative games much easier to analyze. Some game theorists (such as Nash) also have a methodological reason for concentrating on noncooperative games: These theorists hold that such games are "more basic" than cooperative games and that the appropriate way to solve a cooperative game is first to transform it into a noncooperative game. However, these views are not universally held (see Osborne and Rubinstein 1994; Binmore 1992).

A *solution* of a game is a specification of the outcomes that may be expected to occur when the game is played by rational agents. Two widely used techniques for solving noncooperative games are *dominance arguments* and *equilibrium analysis.* The goal of each of these approaches is to identify, for each player, a best-response strategy to the anticipated play of all other players. Given a strategy profile σ, the strategy σ_i is a best-response for player i if $u_i(\sigma_{-i}, \sigma_i) \geq u_i(\sigma_{-i}, \sigma_j)$ for all $\sigma_j \in S_i$, where σ_{-i} denotes the set of strategies in the profile σ for the opponents of player i.

A dominance argument rules out certain strategies for play on the grounds that those strategies are inferior to other alternatives, where an inferior strategy is one that is either weakly or strictly dominated: A strategy σ is *weakly dominated* if there exists another strategy σ' such that the payoff from σ' is never worse than the payoff from σ, and there is at least one instance in which the payoff from σ' exceeds that of σ. A strategy σ is *strongly dominated* if there exists another strategy σ^* such that the payoff given by σ^* always exceeds the payoff given by σ.

Iterated elimination of strongly dominated strategies is a procedure for transforming games into a reduced form. One eliminates the strongly dominated strategies for Player 1, transforming the game G to the game G', and then eliminates the strongly dominated strategies for Player 2 from G' to obtain G'', repeating this procedure until no strongly dominated strategies for any player remain. At the end, one obtains a reduced game G^* with the property that every remaining strategy for every player is a best-response to some possible strategy profile. In addition, the resulting game obtained does not depend on the order or the rate at which strongly dominated strategies are removed. This result does not hold for iterated elimination of weakly dominated strategies. The resulting game G^* obtained by iterated elimination of weakly dominated strategies may depend on the order in which strategies are eliminated, as shown in Figure 3. It is never rational to play a strongly dominated strategy, but there are cases where it is not irrational to play a weakly dominated strategy. Although some game theorists freely apply iterated dominance arguments to reduce the complexity of games, others caution against adopting this as a general approach toward their solution (see Binmore 1992).

Although dominance arguments are useful in analyzing a game, the primary tool of analysis in noncooperative game theory is a Nash equilibrium. A strategy profile σ is a Nash equilibrium if each player's strategy is a best-response to the strategies selected by the rest of the players; alternatively, a Nash equilibrium occurs when no player's expected payoff improves by adopting a different strategy unless another player adopts a different strategy as well. More formally, a strategy profile $\sigma = (\sigma_1, \ldots, \sigma_N)$ is a Nash equilibrium if, for $1 \leq i \leq N$, $u_i(\sigma) \geq u_i(\sigma_{-i}, s_i)$ for all $s_i \in S_i$. The wide acceptance of the Nash equilibrium for solving games derives from the fact that it is the only such concept compatible with the rules of the game, the rationality of the players, and the independent selection of strategies all being common knowledge. (For a discussion of common knowledge, see Lewis 1969.)

If players are restricted to pure strategies, not all games have a Nash equilibrium. The game of Matching Pennies, shown in Figure 4, has no Nash equilibrium when the players are restricted to playing either heads or tails. If players may adopt mixed strategies, then it can be shown that all finite games (that is, games in which each player has only finitely many strategies) have at least one Nash equilibrium (Nash 1950).

	μ_1	μ_2
σ_1	(1,1)	(0, 0)
σ_2	(1, 1)	(2, 1)
σ_3	(0, 0)	(2, 1)

Fig. 3. A game in which order matters for the iterated elimination of weakly dominated strategies.

	Heads	Tails
Heads	(1, −1)	(−1, 1)
Tails	(−1, 1)	(1, −1)

Fig. 4. Matching Pennies, a game with no Nash equilibria (in pure strategies).

Refinements of Nash Equilibrium

Although it is generally agreed that a solution to a game must be a strategy profile in a Nash equilibrium, this provides only a necessary, not a sufficient, condition. In general, Nash equilibria lack several desirable properties: They need not be unique, they need not be optimal, and they may allow players to make incredible threats or promises. The game Battle of the Sexes, shown in Figure 5a, has two Nash equilibria, (Boxing, Boxing) and (Ballet, Ballet). The well-known Prisoner's Dilemma, illustrated in Figure 5b, has (Defect, Defect) as its sole Nash equilibria, yet this outcome yields a payoff of 2 to each player, whereas the outcome (Cooperate, Cooperate) yields payoffs of 3. In the game G (Binmore 1992), (rr, LLL) is a Nash equilibrium, but note that this strategy profile requires that Player 2 commit to playing L at node N, an irrational move, as Player 2 would thereby lose the game if that node were reached, whereas Player 2 would win by playing R. Consequently, a number of refinements and extensions to the concept of a Nash equilibrium have been introduced, two of which are discussed below.

Subgame Perfect Equilibrium

Each node v in an extensive game G induces a subgame of G. A subgame is produced by keeping the node v, along with the subtree rooted at v, and deleting the rest of the game. If σ is a Nash equilibrium of G, it need not be true that σ is a Nash equilibrium for every subgame of G as well. Selten (1965) introduced a refinement of the Nash equilibrium concept known as a *subgame perfect* equilibrium, which requires that a strategy profile σ be a Nash equilbrium for every subgame as well. It has been shown that every finite extensive game of perfect information has at least one subgame perfect equilibrium. Since every subgame perfect equilibrium is also a Nash equilibrium, subgame perfection counts as a refinement of the concept of a Nash equilibrium, because it often eliminates Nash equilibria that are unlikely to be adopted by rational players, such as the strategy profile (rr, LLL) in the Prisoner's Dilemma game of Figure 5b.

Correlated Equilibrium

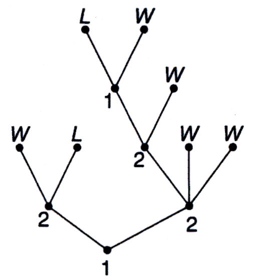
The definition of a Nash equilibrium assumes that the selection of strategies by players occurs independently. Aumann (1974 and 1987) defined a notion of *correlated* equilibrium for noncooperative games. By correlating on shared information about the state of the world (although the information need not be the same for all the players), it is possible for players to arrive at an equilibrium that is self-enforcing in the sense that no player would have reason to deviate from equilibrium play. The fact that correlated equilibria are self-enforcing is significant because it means that adhering to a correlated equilibrium does not require the existence of a binding agreement among the players. In many cases, adopting a strategy profile in correlated equilibrium rewards each player with a higher expected payoff than she could receive in the absence of correlation. For example, consider the game of Battle of the Sexes from Figure 5a and suppose that the players have shared information about the result of a toss of a fair coin. If both players (independently) adopt the strategy of going to a boxing match whenever the coin turns up heads and going to

	Boxing	Ballet
Boxing	(2, 1)	(0, 0)
Ballet	(0, 0)	(1, 2)

(a) Battle of the sexes

	Cooperate	Defect
Cooperate	(3, 3)	(1, 4)
Defect	(4, 1)	(2, 2)

(b) The prisoner's dilemma



(c) The game G

Fig. 5. Games with multiple or suboptimal Nash equilibria.

the ballet whenever the coin turns up tails, each player has an expected payoff of $\frac{3}{2}$, a significant improvement upon their expected payoffs in the absence of correlating their strategies. It has been proven that the set of correlated equilibria always contains the set of Nash equilibria and hence is an extension of the concept of a Nash equilibrium.

Cooperative Games

In a cooperative game, players can enter into binding agreements in which they are committed to playing certain strategies. Whereas strategy profiles in noncooperative games need to be self-enforcing (e.g., a Nash equilibrium) in order to be plausible outcomes of play, in cooperative games the binding agreement can be used to bring about any possible outcome. Of the many possible outcomes, how should one be selected?

Nash (1950 and 1953) proposed the following approach to analyzing cooperative games: Although players *may* enter into a binding agreement, they need not. If they choose not to, then there is a noncooperative game in which each player can, adopting the appropriate mixed strategy, be assured of a certain minimum expected payoff; call this outcome the *disagreement point.* The original cooperative game can thus be conceived as a bargaining problem in which players seek to improve their situation by moving away from the disagreement point to a new, more desirable point conferring greater utility. Exactly which point is selected depends upon the particular arbitration scheme used. An arbitration scheme can be thought of as a function mapping the set of possible outcomes to a single outcome—the solution offered by the arbitrator. A cooperative game, then, can be conceived as an extensive form of a noncooperative game where the early stages of the game involve the selection of the disagreement point and the arbitration scheme. This approach, of reducing cooperative games to noncooperative games, is known as the *Nash program.*

Nash argued that a reasonable arbitration scheme for a bargaining problem should satisfy the following four conditions:

- **Pareto optimality:** It is not possible to increase any player's utility without decreasing another player's utility.
- **Independence of irrelevant alternatives:** The selection of the outcome of the bargaining problem should not depend upon alternatives which were not chosen. (One should be aware that Nash's proposed solution is not universally accepted. This axiom is generally viewed as the most controversial.)
- **Symmetry:** If the set of outcomes is symmetric, then the solution point awards the same payoff to all players.
- **Invariance:** Since utility functions are unique only up to a strictly increasing affine transformation, no player should be able to affect the solution point by rescaling his or her utility function.

The fact that there exists a unique outcome satisfying these four conditions was proved by Nash (1950) for the two-person case.

Solution concepts differing from the one suggested by Nash have been defended by Kalai and Smorodinsky (1975), Braithwaite (1954), and Gauthier (1986). The Kalai-Smorodinsky solution has a natural geometric construction that illustrates the underlying intuitions. Define the "Utopia point" as the outcome awarding each player the maximum amount of utility possible for the game under consideration. In all cases of interest, the Utopia point lies outside the set of feasible solutions. Draw a line l connecting the disagreement point to the Utopia point. The point of intersection between l and the Pareto frontier is the Kalai-Smorodinsky solution. That is, the Kalai-Smorodinsky solution is the point arrived at when each player makes "appropriate" relative concessions from the Utopia point. The solution point identified by the Kalai-Smorodinsky solution is often not the same point as that identified by the Nash axioms.

Evolutionary Game Theory

Evolutionary game theory originated as an application of game theory to biology, arising from the realization that frequency-dependent fitness introduces a strategic aspect into evolution. Evolutionary game theory has since become an object of interest to economists in part because the rationality assumptions underlying it are more appropriate for modeling strategic deliberation by real humans, who are only boundedly rational, as opposed to the perfectly rational agents modeled by traditional game theory. In addition, evolutionary game theory provides a way of modeling the dynamics of strategic interaction in a way not possible with the traditional theory of games. Recall that the only way to model the temporal aspect of a game is to use the extensive form of representation. However, methods of analyzing extensive games typically proceed by envisioning that players select

	Rock	Paper	Scissors
Rock	(1, 1)	(0, 2)	(2, 0)
Paper	(2, 0)	(1, 1)	(0, 2)
Scissors	(0, 2)	(2, 0)	(1, 1)

Fig. 6. The game of Rock–Paper–Scissors.

a strategy at the beginning of the game that specifies their course of action at each choice point, which really does not model the dynamical aspect of the game.

The primary equilibrium concept in evolutionary game theory is that of an evolutionarily stable strategy (see Maynard Smith 1982). A strategy is *evolutionarily stable* if when almost every member of the population follows it, no individual who adopts a novel strategy can successfully invade. If σ is evolutionarily stable, the fitness of an individual following σ must be greater than the fitness of an individual following μ (otherwise the individual following μ would be able to invade, and so σ would not be evolutionarily stable). Let $F(s_1, s_2)$ denote the change in fitness for an individual who plays the strategy s_1 against an opponent playing the strategy s_2. Then σ is evolutionarily stable if and only if:

$$F(\sigma, \sigma) > F(\mu, \sigma)$$

or

$$F(\sigma, \sigma) = F(\mu, \sigma) \text{ and } F(\sigma, \mu) > F(\mu, \mu).$$

If a strategy is evolutionarily stable, it must be a best reply against itself, for, if not, a mutant strategy would be able to invade. This means that all evolutionarily stable strategies are Nash equilbria when played against themselves. However, not all games have evolutionarily stable strategies, and not all Nash equilibria are evolutionarily stable. The game of Rock–Paper–Scissors, shown in Figure 6, has a unique Nash equilibrium in mixed strategies where each individual plays Rock, Paper, or Scissors with probability $\frac{1}{3}$, but no evolutionarily stable strategy.

J. McKenzie Alexander

References

Aumann, R. (1974), "Subjectivity and Correlation in Randomized Strategies," *Journal of Mathematical Economics* 1: 67–96.

——— (1987), "Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrika* 55: 1–18.

Binmore, K. (1992), *Fun and Games*. Lexington, MA: D. C. Heath and Company.

Borel, É. ([1921] 1953), "The Theory of Play and Integral Equations with Skew Symmetric Kernels" (translated by L. J. Savage), *Econometrika* 21: 97–100. Originally published as "La théorie du jeu et les équations, intégrales à noyau symmétrique gache," *Comptes Renduc de l'Académie des Sciences* 173: 1304–1308.

Braithwaite, R. B. (1954), *Theory of Games As a Tool for the Moral Philosopher*. Cambridge: Cambridge University Press.

Gauthier, David (1986), *Morals by Agreement*. Oxford: Oxford University Press.

Harper, W. (1988). "Causal Decision Theory and Game Theory: A Classic Argument for Equilibrium Solutions, a Defense of Weak Equilibria, and a New Problem for the Normal Form Representation," in W. Harper and B. Skyrms (eds.), *Causation in Decision, Belief Change, and Statistics II*. Dordrecht, Netherlands: Kluwer.

Kalai, E., and M. Smorodinsky (1975), "Other Solutions to Nash's Bargaining Problem," *Econometrika* 43: 513–518.

Kreps, D. M. (1990), *Game Theory and Economic Modelling*. Oxford: Oxford University Press.

Lewis, David (1969), *Convention: A Philosophical Study*. Oxford: Basil Blackwell.

Luce, R. D., and H. Raiffa (1957), *Games and Decisions*. New York: Wiley.

Maynard Smith, J. (1982), *Evolution and the Theory of Games*. Cambridge: Cambridge University Press.

Myerson, R. B. (1991), *Game Theory: Analysis of Conflict*. Cambridge, MA: Harvard University Press.

Nash, J. (1950), "Equilibrium Points in N-person Games," *Proceedings of the National Academy of Sciences of the United States of America* 36: 48–49.

——— (1953), "Two-Person Cooperative Games," *Econometrica* 21: 128–140.

Osborne, M. J., and A. Rubinstein (1994), *A Course in Game Theory*. Cambridge, MA: MIT Press.

Samuelson, L. (1998), *Evolutionary Games and Equilibrium Selection*. Cambridge, MA: MIT Press.

Selten, R. (1965), "Spieltheoretische Behandlung eines Oligopolmodells mit Nachfragetragheit," *Zeitschrift für die gesampte Staatswissenschaft* 121: 301–324.

von Neumann, John ([1928] 1959), "On the Theory of Games of Strategy" (translated by Sonya Bargmann), in A. W. Tucker and R. D. Luce (eds.), *Contributions to the Theory of Games*, vol. 4 (*Annals of Mathematics Studies* 40). Princeton, NJ: Princeton University Press, 13–43. Originally published as "Zur Theorie der Geselleschaftspiele," *Mathematische Annalen* 100: 295–320.

von Neumann, John, and Oskar Morgenstern (1944), *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.

Zermelo, E. (1913), "Über eine Anwendung der Mengenlehre auf die Theorie des Schachspiels," in *Proceedings of the Fifth International Congress of Mathematicians* 2: 501–504.