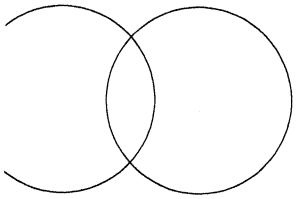


COOPERATION AND CONFLICT

In This Chapter You Will

- Analyze the problems of collective action in light of the Prisoner's Dilemma.
- See how people can learn to cooperate by using the "tit for tat" criterion in repeated interactions.
- Think about why cooperation is more likely to hold up in long-term relationships.
- Reflect on the advantages of and drawbacks to unilateral cooperation for the provision of environmental public goods.

Climate change has become a broad concern, leading many people to call for action to stop the planet from overheating. Yet some skepticism persists, and few governments are willing to deal with the problem by themselves. The Kyoto Protocol to put a limit on emission of greenhouse gases was formally accepted by thirty-six developed countries for the period 2005–2012. However, the United States and Australia initially refused to sign. Even some of the protocol signatories, and many developing countries, including most prominently China, continue to grow and pollute as much as they like. The holders of the global warming thesis warn that if the international response did not substantiate cooperation, the world might be condemned to catastrophic events.

However, recently, most countries seem to have found some incentive for avoiding being sanctioned for their misconduct. The Asia-Pacific Partnership on Clean Development and Climate has attained new cooperation on development and technology transfer, enabling a reduction in gas emissions. The United Nations has held talks with virtually all countries in the world to replace the Kyoto Protocol after its deadline. If things get bad enough, then cooperation may flourish.

This kind of problematic decisions on whether to cooperate or not can be analyzed with the help of game theory. In this chapter we will see more formally how, when

cooperation is desirable but difficult, it can emerge and hold up for the provision of public goods such as a clean environment.

The Prisoner's Dilemma

Game theory studies human decisions in situations in which one's decision depends on expectations as to what others will do. They are not like individual decisions to purchase one or another good in a supermarket. They are more comparable to situations such as trying to build an international alliance among different states, such as the one regarding climate change just mentioned; negotiating between two political parties to form an electoral coalition; bargaining between a seller and a buyer to reach an agreeable price; or even playing cards with a few friends. Game theory has many applications in politics, as well as in economics, sociology, psychology, biology, and ethics.

In a strategic situation requiring collective action, each actor must decide whether to **cooperate, that is, to participate in the common action for the common benefit**, or **to compete, that is, to seek his own benefit**. When making his decision, each actor must consider how others may respond to that decision, because others' cooperation or defection can change the collective outcome. It is not isolated individual decisions but, rather, interactions among several actors that produce collective outcomes.

In game theory, different types of interactions are distinguished, especially the following:

- **Coordination games**. These imply easy cooperation among people with strongly shared interests to produce common benefit.
- **Conflict or zero-sum games**. In contrast to the previous coordination games, these include strong competition in which **gains for some participants imply losses for the others**.
- **Non-zero-sum games**. Falling in between coordination and conflict games, these are games in which the outcome is undetermined, whether in favor of cooperation or conflict, although **mutual cooperation can produce gains for all participants**.

COOPERATE AND DEFECT

A non-zero-sum game that is particularly important to the science of politics is the **Prisoner's Dilemma**. This game has **an inefficient outcome in which nobody cooperates**, which provides insight into the problems of collective action. As we have seen, on many occasions people fail to cooperate with others even if cooperation would produce a better collective outcome for all the participants. The lessons from the Prisoner's Dilemma can be applied to any group or community facing a cooperation problem among its members.

The metaphorical story of the Prisoner's Dilemma is about two burglars who have been arrested by the police as the suspects of a crime. Let us call them John and Oskar. The police chief questions John and Oskar in separate rooms and offers each of them the following deal:

"If you confess to the robbery and report your fellow-mate, your declaration will be taken as proof, and he will be condemned to ten years in prison, but as a prize for

your declaration you will go free. If you both confess to the crime by denouncing each other, you both will be declared guilty and each of you will get a sentence of five years. If, however, neither of you confesses, we can lock you up for one year for illegal possession of arms."

Each prisoner has two strategies: **C, cooperate** (remain silent), and **D, defect** (denounce the other). Table 3.1 shows the outcomes associated with each prisoner's decision.

Note that John and Oskar are in separate cells and cannot communicate with each other but can only speculate about the other's likely decision. What would you expect them to do? Would they cooperate and remain silent, or denounce each other? The collective outcome—that is, the sentence for the two prisoners—depends on the strategy each chooses and the strategy chosen by the other.

Consider John's decision first. He does not know what Oskar is going to do. If he presumes that Oskar will cooperate and remain silent, John's best strategy is to denounce him, since then he will go free rather than spending one year in prison. If, on the contrary, Oskar were to denounce John, John's best strategy is still to denounce Oskar, since then he will spend five years in prison rather than ten. So, regardless of what Oskar does, John's best strategy is to denounce him.

The possible strategic choices are represented in Table 3.2. John's decisions can be approached from the left of the table, in two different rows, while Oskar is placed at

TABLE 3.1 Individual Preferences

In the game between two criminals suspected of committing a crime, each can be free by accusing the other.

		COLLECTIVE OUTCOME	
JOHN'S DECISION	OSKAR'S DECISION	JOHN	OSKAR
D defect	C cooperate	Free	10 years
C cooperate	C cooperate	1 year	1 year
D defect	D defect	5 years	5 years
C cooperate	D defect	10 years	Free

TABLE 3.2 The Prisoner's Dilemma

The sentence to be received by each prisoner depends both on his decision either to cooperate with or denounce the other fellow and on the decision made by the other prisoner.

		OSKAR	
		COOPERATE C	DEFECT D
JOHN	COOPERATE C (*remain silent*)	1 year for each	John gets 10 years Oskar goes free
	DEFECT D (*denounce*)	John goes free Oskar gets 10 years	5 years for each

the top of the table, over two columns of possible decisions. With the assumption that Oskar cooperates and remains silent—that is, if he chooses C and places the possible outcomes at the left column—then John's best strategy is to denounce, D, represented by the last row. The outcome is located in the lower-left cell: John goes free while Oskar gets ten years. If, on the contrary, Oskar denounces John and chooses D, in the right column, then John's best strategy is still to denounce Oskar, which leads to five years in prison for each of the two, as represented by the lower-right cell.

In the language of game theory, we say that John has **a dominant strategy, that is, a decision resulting in a better outcome for him regardless of the other actor's decision**. In this case, denouncing is a dominant strategy for John because he will get less time in prison if he denounces Oskar, regardless of whether Oskar remains silent or denounces John.

Now consider Oskar's decision. Since the strategic situation is symmetric for each prisoner, he faces the same choices as John, and he can reason in much the same way. Check the table again to see this. Regardless of what John does, Oskar will get less time in prison if he denounces him. So denouncing is also a dominant strategy for Oskar.

Thus the two actors may follow their dominant strategies, denouncing each other, and both will spend five years in prison. This outcome is represented by the lower-right cell of Table 3.2, as the result of the two prisoners' decisions to choose strategy D. Certainly this is a very bad, inefficient outcome. The two would have been better off—that is, they would have gotten less years in prison—had they both remained silent. By cooperating and not denouncing each other—that is, by choosing strategy C—they would have spent only one year in prison for possessing a gun, as represented by the upper-left cell of Table 3.2. But by each pursuing his individual interest in isolation from the other, the two prisoners choose a mutual betrayal on the basis of the fear that the other fellow will also choose to denounce. They reach a collective outcome that is worse for each of them.

The outcome produced by two dominant strategies is said to be an **"equilibrium,"** in the sense that it is **stable**. Consider the possibility of John withdrawing his denunciation and denying his previous finger-pointing. He would be declared the only one guilty and would have to bear the full penalty of ten years in prison. In Table 3.2, he would move the outcome from the lower-right cell to the upper-right one. Nor does Oskar have an incentive for unilaterally modifying his strategy of denouncing his partner. If neither actor were to change his strategy unilaterally, given the chance, the outcome would be considered stable, according to the concept of equilibrium coined by Nobel Prize-winning mathematician John F. Nash. **The "Nash-equilibrium" is an outcome from which no actor has incentives to move away by changing his strategy unilaterally.**

The inefficient outcome of the Prisoner's Dilemma is a representation of the free-rider problem in collective action. Each individual actor pursues his best interest by pretending to use public goods without contributing to their provision, but in the end nobody cooperates, and they cannot satisfy their common interest.

Note, however, these key assumptions in the story of the prisoner's dilemma: the actors cannot communicate, and their choices are not retractable. Once an accusation is made, one's fellow prisoner becomes irremediably guilty. It is not possible to retract

the accusatory information given, just as it is not possible to deny proven facts. The non-retractability of the choices unavoidably produces an undesirable outcome. Yet this assumption may not fit some real-life situations of human interactions whose formal structure can be represented with this game, especially if they involve a low number of people who can easily communicate with each other.

In certain settings, the game may not end once the actors have made their simultaneous choices, because, on the basis of communication and the knowledge of the other actor's decision, each actor still has the chance of retracting his initial choice. A model of bargaining or negotiation between two actors, for instance, may include the possibility for one to make an offer. Yet if the other does not accept it, the initial choice does not irreversibly damage him. The bargain is not completed, and the former can listen to the latter's counteroffer or propose a new one. This can be applied to interactions between two political parties or two governments, and between a seller and a buyer, and in many other situations. The game ends only when every actor receives a response to his choice from the other actor. Then an actor can start offering conditional cooperation and expect a positive reply from the other actor if he has the possibility to change his strategy. This kind of cooperative agreement to the actors' mutual benefit is more likely to be reached in a small group with direct interactions. In contrast, in a large group with numerous actors, the chances of obtaining positive responses from the other actors and of dissuading them from defecting may still be slim.

OTHER POLITICAL DILEMMAS

As mentioned, the Prisoner's Dilemma can be used to understand the problems of collective action for the provision of public goods, including, as in our initial example, saving the atmosphere. As the players in the game have dominant strategies leading to an inefficient collective outcome, so the participants in collective action may be tempted to free-ride on the others and produce insufficient provision of public goods. In a number of cases, however, closer interaction between actors on problems involving high stakes can foster conditional cooperation and agreeable, more satisfactory outcomes. The same logic applies to many other situations, up to the point that the Prisoner's Dilemma can be considered one of the better-studied models of social interactions. Consider the following examples.

Community. What certain classical authors called "the state of nature" can be understood as a multi-actor Prisoner's Dilemma in which conflict is pervasive. If human interactions are unconstrained, anybody with the advantage of surprise can try to impose his will on the others. But if all do this, then people may find themselves living in a state of chaos in which, in Thomas Hobbes's famous words, life tends to be "solitary, poor, nasty, brutish, and short." In such an environment it is not reasonable to risk unilateral cooperation, while cooperation within groups is precarious. However, human beings can do it better. People can agree on creating a government equipped with coercion tools to enforce rules mandating those actions that individuals find beneficial to all. The government may apply sanctions against "defectors"—that is, violators of mutually beneficial rules of conduct—discourage free-riding on public goods, and craft incentives for cooperation. People can rationally accept conditional consent. By an agreed "social contract," the efficient outcome of civilization or "commonwealth," in which each can live in peace and security, can be attained.

BOX 3.1

GOING FOR THE BLONDE

One of the most influential contributors to game theory is John F. Nash, who was awarded the Nobel Prize in Economics in 1994 (probably the Nobel winner with the fewest pages published in his life!). In three articles, he basically introduced two concepts. The so-called Nash equilibrium is the outcome of a game in which no player can improve his results by unilaterally changing his strategy, as explained in this chapter.

Nash also introduced a normative concept, the "Nash solution," for bargaining problems in which two or more actors try to reach an agreement on how to divide a good. The classical utilitarian criterion holds that the best social solution is the one that maximizes the sum of the actors' utilities. This criterion inspires, for instance, the evaluation of the welfare of a country by its average per capita income or by the general level of citizens' political satisfaction. In contrast, the Nash solution is the one that maximizes the *product* of the actors' utilities.

An example of the Nash solution is given in the Oscar-winning movie *A Beautiful Mind* (based on the book of the same title by Sylvia Nasar) about the life of John Nash. A group of four students, including Nash, are spending some time in a bar in Princeton when five girls, including a stunning blonde, enter the room. Nash says, "If we all go for the blonde, we block each other and not a single one of us is gonna get her." He suggests that "no one goes for the blonde," and that the four boys instead pair up with the other girls—to the bemusement of his fellow students.

A reasonable interpretation is the following. If only one person obtains the maximum prize, say with value 10, and the other three get nothing, the sum of the four actors' utilities can be relatively high in comparison with an alternative outcome in which each of the four actors gets a lower value, say of only 2 (since $10 + 0 + 0 + 0 = 10 > 2 + 2 + 2 + 2 = 8$). But the product of utilities in the first outcome is very bad (actually it is zero, since three actors get nothing and can be extremely frustrated), while the product of the latter is higher ($2 * 2 * 2 * 2 = 16$).

Generally, in comparison with the classical utilitarian "Bentham sum" solution, the "Nash product" solution favors more egalitarian distributions. For example, under the sum criterion, a distribution of values among three actors such as 3, 2, 1, is as good as the distribution 2, 2, 2 (since both imply a sum of 6 units of social utility). But under the product criterion, the former distribution (whose product is $3 * 2 * 1 = 6$) is worse than the latter one (whose product is $2 * 2 * 2 = 8$).

Democratization. In situations of institutional regime crisis in which authoritarian rulers cannot go on as they have been accustomed to, actors with opposite political regime preferences can generate violent conflict or a civil war in which each side may fight to eliminate the other, as represented by the inefficient outcome in the Prisoner's Dilemma. Eventually, one of the sides can become a single, absolute winner, but choosing confrontation with uncertain outcome also entails the risk of becoming an absolute loser, and costs of significant destruction on both sides. In contrast, by anticipating the foreseeable consequences of their choices, either the rulers or the opposition leaders can offer conditional, retractable cooperation. Negotiations can lead to a provisional compromise, including the calling of a multiparty election not securing an absolute winner, which may open further developments in favor of either of the actors involved, as has happened in so many cases of democratization in different parts of the world since the last quarter of the twentieth century.

Deterrence. International relations during the cold war between the United States and the Soviet Union were much like the Prisoner's Dilemma. If the Soviets chose to build new weapons, the United States did the same, and vice versa, which

CASE 3.1 PRISONER'S DILEMMA AT THE OPERA

The plot of the opera *Tosca*, written by Giacomo Puccini (1858–1924), is a case in point about the troubles that can be created by a Prisoner's Dilemma–type of situation. In the story, a new republic has been established in Rome under the influence of the French Revolution. Scarpia, the reactionary chief of police, captures the rebel Mario and offers him a deal: his life in return for his lover Tosca's physical surrender. Tosca has to concede, and Scarpia, to keep up appearances, gives order for a mock execution of Mario with blank cartridges. However, Tosca manages to avoid intimacy with Scarpia by stabbing a knife into his chest. Then she runs to the firing squad ceremony. After the soldiers shoot, she finds Mario is actually dead. Scarpia has double-crossed her. The interaction can be represented as follows:

		POLICEMAN SCARPIA	
		MOCK EXECUTION	REAL EXECUTION
TOSCA	CONCEDE	Mario alive Tosca yielded	Mario dead Tosca yielded
	RESIST	Mario alive Tosca unharmed	Mario dead Tosca unharmed

Note that the initial compromise, as represented by the upper-left cell, is not an equilibrium, since the two actors have incentives to modify their decision unilaterally. Tosca has a dominant strategy: to resist Scarpia's advances, regardless of Scarpia's choice of either leaving Mario alive or executing him, while Scarpia also has a dominant strategy of operating a real execution whether Tosca concedes or resists. The inefficient outcome of the Prisoner's Dilemma is represented by the lower-right cell, although the two actors would have been better off in the upper-left cell.

triggered an arms race that put both countries at risk. For each power, arming was a dominant strategy. The unrelenting logic of self-interest drove the two powers toward a conflict outcome that was worse for each of them than living in a world safe from the arms threat. The "balance of terror" without actual direct war was sustainable because, with nuclear weapons and the possibility of total destruction, the stakes were so high. Each power was able to prevent the other from ceding to the temptation of unilateral surprise attack, since each could threaten the other with massive retaliation. But the strength of mutual deterrence indirectly provoked lots of non-nuclear, limited war outbursts from local conflicts throughout the world. The United States and the Soviet Union attempted to move away from the inefficient outcome of the game—that is, to reduce and control arms through negotiations and agreements. But the arms race game stopped only with the dissolution of one of the players.

Quick Quiz

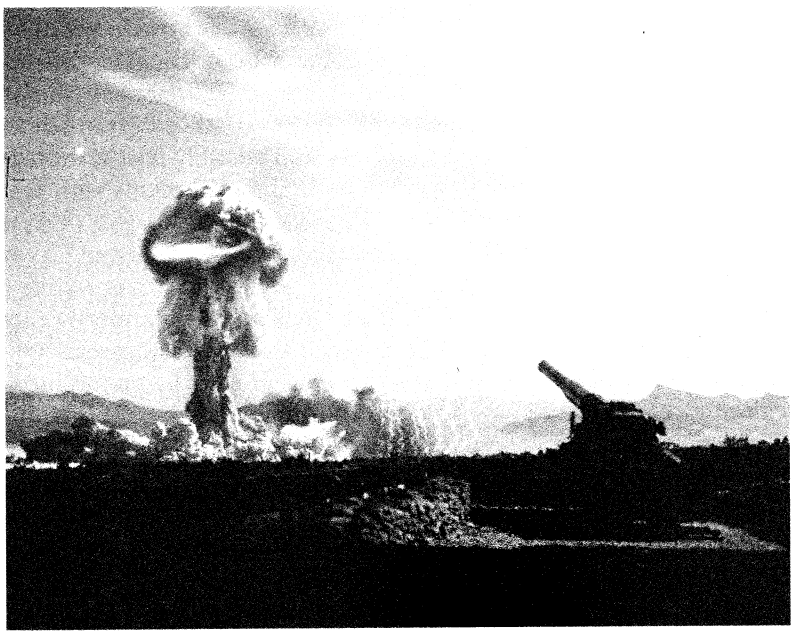
- What's an equilibrium?
- What's a dominant strategy?

SOURCE 3.1 Theory of Games

The origins of game theory were related this way by one of its inventors:

The Nazis took over in Vienna in March 1938. I [*Oskar Morgenstern*] was dismissed as "politically unbearable" from the University as well as from my Institute... While in the United States, I received a number of calls from various American universities to join their faculties... The principal reason for my wanting to go to Princeton was the possibility that I might become acquainted with [*John*] von Neumann [*also in exile from communism in Hungary and Nazism in Germany*] and the hope that this would be a great stimulus for my future work. Von Neumann and I met soon after the University opened... We were already in the midst of war, not an ideal time for work of our kind!... I recall vividly how Johnny rose from our table when we had set down our expected utility axioms and called out in astonishment: "Ja hat denn das niemand gesehen?" ("But didn't anyone see that?")... Clearly, we were convinced that it represented first of all a fundamental break with conventional economics: we demonstrated that one is not confronted with ordinary maximum or minimum problems (no matter what side conditions!), but with conceptually different situations. Though this becomes intuitively quite easily accessible for ordinary exchange, for monopoly, oligopoly, etc., the phenomenon is all-pervading... Thus the scope of the book extends far beyond economics, reaching into political science, sociology, etc.

From Oskar Morgenstern, "The Collaboration Between Oskar Morgenstern and John von Neumann on the Theory of Games" (1976)



Arms races, such as during the Cold War between the United States and the Soviet Union, can be analyzed with the Prisoner's Dilemma Game.

The Evolution of Cooperation

The decision to cooperate or defect may not be the same if you are involved in a single occasion to interact with unknown people or if you are going to continue to interact with the same group of individuals for a long time. If the "game" consists of only one play, many people may think that a bird in the hand is worth two in the bush. This is the kind of decision some people tend to make in a Prisoner's Dilemma situation, such as a street market in which any passerby can bargain with anonymous strangers they are not likely to meet again.

The situation is different if people are going to engage in repeated plays. Then it may make sense to try to cooperate in order to receive others' cooperation in the future. A community or institutional setting in which everybody can expect to keep interacting with the same people regularly for some time may include a household or a neighborhood, the workplace, a mall or a school, a professional organization or a political party, and the city, state, empire or world in which one is aware of living in and where one intends to stay. Reasonable behavior for repeated interactions with other individuals with some common interest can involve conditional cooperation and a positive response to the others' behavior. The possibility of cooperation in repeated plays of Prisoner's Dilemma–type situations can derive from a decision criterion called "Tit for Tat." In the long term, cooperation may spread and become the prevailing way of conduct, as we discuss in the following pages.

TIT FOR TAT

If you have to decide whether to cooperate or defect in a series of interactions, it is convenient to have a criterion for making each decision. A criterion for choosing a strategy in every interaction is called a "meta-strategy." Let us consider different possibilities for interactions that have the logical structure of the Prisoner's Dilemma or of collective action in general. First, you may choose, for instance, "always defect." This is indeed the safest meta-strategy, in which no one can possibly take advantage of you. But after a while you may get stuck in a rut of mutual punishments with your fellow "prisoners," which is likely not to produce any collective good.

A second meta-strategy may be just the opposite: "always cooperate." It works well if everybody else applies the same criterion. But it may get very bad results if others defect, since the cooperative actor can become a "sucker." It is like always turning the other cheek—and getting slapped. As philosopher Jean-Jacques Rousseau noted many years ago, "for the society to be peaceable and for harmony to be maintained," all the citizens *without exception* would have to be good cooperators; but "if by ill hap there should be a *single* self-seeker or hypocrite, he would certainly get the better of his pious compatriots... servitude and dependence."

With any of the two meta-strategies mentioned—always defect or always cooperate—you would always be running on autopilot, not taking into account what other people were doing. If you do take their behavior into account, you can improve the results greatly. Specifically, if after a few interactions you realize that the other people are always cooperating, then you can defect and gain advantage. If, on the contrary, the other people are always defecting, no matter how cooperative you are, then you should always defect, too, in order to avoid being slapped again and again and becoming a sucker. But in general you can expect that the other people will

respond to your actions, too. What would be the best way for you to act if there are real interactions, that is, responsive actions by other people?

The Tit-for-Tat meta-strategy, or criterion of conduct, for repeated plays of the Prisoner's Dilemma can produce sustained cooperation. It is based on the idea of responding to the others' behavior or **doing unto others as they do unto you**, as initially identified by political scientist Robert Axelrod. As a recommendation for conduct, it includes the following elements:

- **Start nice.** Don't be the first to defect. In some real quarrels it is not even clear who started it; both sides can claim that the other started slapping and that they are just responding. This happens in situations ranging from children at play to neighboring states in conflict. The solution is, of course, to start again by making clear the initial choice of each party. Let us, therefore, start with cooperation, see what happens—that is, how the other people respond—and then do the same as they do.
- **Retaliate.** Cooperate if the others cooperate; defect if the others defect. In a series of plays, your reply should be given at the immediate next play of the game.
- **Forgive.** Don't be resentful. After you have replied to the others' defection with your own defection, be willing to start conditional cooperation again. Your negative response to the previous defection of the other actor should be a lesson, sending a message to him that he should learn to cooperate.
- **Keep it simple.** Basically, do unto others as they do unto you. Your best lemma could be: "If you cooperate, I do; if you don't, I don't." By doing this, the efficient outcome in the Prisoner's Dilemma–type of interaction can be achieved again and again. Cooperation will be met with cooperation and will feed on itself. In the same sense, the free-rider problem in collective action can be solved. People can indeed organize themselves in a stable manner and contribute to the provision of collective goods in which they are genuinely interested.

INDEFINITE PLAYS

A more pessimistic view holds that the incentive to cooperate in a series of Prisoner's Dilemma–type interactions may vanish if the participants know when the "game" is going to finish—that is, when they will cease to be members of the same community sharing some common interest. The last play of an iterated Prisoner's Dilemma is like a single play, because there will be no further opportunity to reply to any defection. Everybody will have the dominant strategy to defect—as in the one-shot interaction between John and Oskar from their cells.

Actually the problem may be even worse. If you know when the end of the game will be, then you can expect that the other actor will defect at the last play—as will you. Then you discount that play. Considering now the next-to-last play, since you have discarded any response in the last play, the next-to-last now appears as the last play on which to make a decision. Again, this will be like a single play, in which you will also defect, as will the other actor. You and every actor can reason in the same way at every play, and thus decide to defect at every "previous" play up to the beginning. By so-called **"backward induction," the repeated interaction can lead to mutual defection when the end is known**, The iteration of Prisoner's Dilemma–type interactions,

if the end of the iteration is known in advance, is like a single play of the game. For all the actors, therefore, the dominant strategy is to defect in all plays of the game.

Cooperation by means of tit for tat can be achievable if the number of interactions is infinite, which is not likely to occur in this world, or if they are indefinite—that is, if the actors don't know when their interactions will finish. If nobody knows how many plays lie ahead, there is no "last" play to reason backward from. In their ignorance of the future, people will keep cooperating. Those who know, however, can take advantage at the last play. In more practical terms, we can expect that cooperation is more likely to hold up the greater the uncertainty about the length of the game and the higher the number of plays or interactions. The more repeated the cooperation strategy, the wider the spread both of information and of the actors' reputation regarding their cooperativeness.

According to this insight, we can observe that cooperation, collective action, and joint organization are indeed more intense and sustained among certain groups of people interacting for long periods of indefinite length. Cooperation should be higher, for example, among members of a condominium complex rather than among motel clients; among town residents rather than among tourists or occasional visitors; among fixed employees rather than among temporarily unemployed people expecting to find another job soon; among civil servants rather than among seasonal workers; among store owners in a commercial mall rather than among sporadic vendors in a street market; among practitioners of professions requiring costly training or implying low opportunity costs, such as miners or physicians, who are likely to stay in the job, rather than among amateurs or aficionados; among students enrolled in three- or four-year programs rather than among summer-course attendees; and among citizens in countries with a sedentary population rather than in those in which many people are likely to emigrate.

FAVORABLE CONTEXTS

Criteria of behavior such as Tit for Tat promise satisfactory results only in the long term and in average values. But no meta-strategy guarantees an efficient outcome at every play of the game. Particular results in specific circumstances may depend on the social context, that is, on what other people are doing in the particular group in which you are participating. Actually you can choose your partners—that is, you can decide to interact preferably with people with whom you can establish beneficial relations. Once you are in a particular group or community, some criteria of conduct may prove to be better than others. In general it is advisable to follow the dictum "When in Rome, do as the Romans do," although this may require some adaptation.

Self-sustained Cooperation. Specifically, if in a group everybody applies Tit for Tat, then Tit for Tat is the best strategy for each individual. Cooperation is universal and permanent. In such a context, even the meta-strategy "always cooperate" can survive and obtain the same result, since people using Tit for Tat reply to cooperation with cooperation.

In certain contexts, slightly different meta-strategies might produce even better results. **"Tit for Two Tat,"** for instance—which implies not responding with defection to the other actor's first defection but only to the second one, or, in other words,

being more forgiving—may help prevent the quick diffusion of defective behavior in communities in which most people are prone to behaving cooperatively.

Even in an adverse context in which people using the "always defect" criterion abound, a set of people applying Tit for Tat may obtain some success. After the permanent defectors have proven their advantage, the population of exploited "suckers" cooperating unilaterally tends to disappear. Then the defectors may find themselves with nobody to exploit, thus mutually defecting with everybody, which produces mediocre results. In the long term, people with the Tit for Tat criterion can obtain better results than defectors by responding with defection to the defectors and developing numerous cooperative interactions among themselves. This may make Tit for Tat advantageous and eventually the most common strategy in the group.

Trust. If people reciprocate with cooperation to cooperation in repeated interactions, they can build a good reputation for themselves that may move other people to cooperate with them. Feelings of trust may emerge among people having information about others' past action and among new participants obtaining regular positive retribution for their conduct. In the mid- or long term, increasing and sustained cooperation among members of a community may induce them to construct institutional environments that limit individual competition and tend to homogenize the population. Internal sanctions against defectors can go together with the promotion of values such as honesty and empathy for others' distress, thus reinforcing social cooperativeness.

Indeed a sense of reciprocity and a capacity for empathy seem to have developed over time among human beings. In fact, there are numerous groups and communities in which broad conditional cooperation prevails, although with some degree of diversity, including both good unconditional cooperatives and evil defectors—after all, there are always saints and sinners in the world.

Intergroup Conflict. The development of cooperation among members of a community or interest group requires some mutual commitment to stay within the contours of the group. If, conversely, people living within the same institutional setting consider themselves to belong to two different groups with opposite goals, asymmetric relations can develop. The difference between groups can be based on family or tribal traditions, contrary economic interests, adversarial preferences for the location of public goods, or alternative ethnic allegiances such as language, race, or religion.

People can then apply "discriminatory Tit for Tat." When interacting with members of the same group, as it is perceived by the individuals themselves, they can start cooperating and develop the afore-mentioned kind of evolution favoring collectively efficient outcomes to their mutual benefit. People may be more cooperative toward others the more closely related they are. But when two individuals from clearly different groups interact, they can defect. This arrangement can be stable. In this context, an individual who tries to use the nondiscriminatory criterion of Tit for Tat will be worse off because his cooperative strategy with members of the other group will be regularly responded to with defection. If this is observed, most people may tend to

start any interaction with members of the other group with defection, just to prevent being the sucker from the beginning.

In this type of situation, the larger group may obtain relatively better results than the smaller group. Even if each individual interacts randomly with members of any of the groups, the members of the larger group interact more often with other members of the same group, thus having more opportunities to develop cooperative strategies leading to efficient outcomes. The members of the smaller group, by contrast, are likely to have more interactions with members of the larger group, thus receiving higher proportions of inefficient results produced by mutual defection.

In the long term, members of the small, disadvantaged group may want to split off and create a separate community, whether another family, tribe, company, organization, union, coalition, municipality, state, or alliance. This may be beneficial for members of all groups, since each will then have better opportunities to interact within their own group and to develop cooperation as a consequence of iterated plays.

Quick Quiz

- What's a meta-strategy?
- Define the "Tit-for-Tat" meta-strategy in two sentences.

Other Games of Collective Action

Although the Prisoner's Dilemma can be considered as representing the most basic structure of the problems of collective action for the provision of public goods, other game structures can clarify certain other situations. People can play different "games"—that is, they can develop different strategic interactions, depending on the opportunities, incentives, and constraints supplied by given structures and institutions. To put it in terms of critical political economist Karl Marx, "Men make their own history, but they do not make it as they please... but under circumstances existing already, given and transmitted from the past." Structural conditions do not determine human behavior, but usually shape a distinct set of feasible choices. Human beings can also try to modify certain elements of a given situation, especially rules, institutions, and organizations, in order to have the capacity to develop the interactions in which they are interested.

Game theory can help to model specific problems for the provision of different types of goods and in different settings in which individuals can organize collective action, which we have addressed more informally in the previous two chapters. In the following pages we discuss two well-known games that correspond to different sets of incentives and opportunities: "Chicken" and "Stag Hunt" (curiously, both with animal names). For the Prisoner's Dilemma and for each of these two additional types of situations, different outcomes—that is, different levels of cooperation among members of an interest group—can be expected.

CHICKEN

Besides the Prisoner's Dilemma, another important model for a diverse range of human interactions is "Chicken." It can enlighten, among many other situations,

some challenges of the global warming problem mentioned at the beginning of this chapter. The origin of the model's name is a real Saturday night game played by some teenagers in American cities in the 1950s and popularized across the world by movies such as *Rebel Without a Cause*. One variant of the game consists of two boys simultaneously driving cars toward the edge of a cliff and either putting on the brakes or jumping out at the last possible moment. The one exiting the car closest to the edge wins, and the other is declared the "chicken."

In this game it is not difficult to identify the preferences of the players, let us call them James and Dean. First, the best thing that can happen to James is to keep driving ahead and jump out and let Dean stop. James wins, and Dean is a chicken. The second best thing is that both boys stop in good time; both are chickens but save their lives. Finally, James's stopping first while Dean continues driving implies that he is a chicken, but it is less bad than if no car stops and both boys fall over the edge of the precipice. Dean has analogous preferences.

Let us call braking the car "cooperating" and driving ahead "defecting." Note the difference between the structure of this game and that of the Prisoner's Dilemma, which is apparently minor but has significant consequences. The actors in the two games switch only their last and next-to-last preferences. In the Prisoner's Dilemma everybody prefers to defect if the other actor defects. Cooperating when the other defects would make one the sucker—the worst possible outcome. In contrast, in Chicken, unilateral cooperation while the other defects—putting on the brakes and saving his life—is less bad than mutual defection, which implies that both players go over the cliff.

What would you expect the outcome of the game to be? Consider James's decision. He does not know what Dean is going to do. If he sees that Dean is stopping, he should drive just a little farther ahead and jump out of his moving car. If, in contrast, James thinks that Dean is going to drive ahead, it is better to stop—better a chicken than dead. In contrast to Prisoner's Dilemma–type situations, here the actors do not have dominant strategies. In one assumption about the other player's decision, it is better to drive ahead, while in the other assumption it is better to stop. Thus there are **two possible outcomes of the game in which some actors cooperate and others defect.** In Chicken, in one outcome James wins and Dean stops; while in the other, James stops and Dean wins. But which of the two possible outcomes will occur is uncertain—that's precisely the allure of the game. Check the players' possible strategic choices in Table 3.3. The two outcomes of the game are located in the upper-right and the lower-left cells.

One way to try to select one of the two possible outcomes of the game may imply pre-commitment by one of the players not to stop. If at the beginning of the challenge James announces out loud before his fans that he will not stop, he may force Dean to stop early (or just quit the game). But any actor's pre-commitment must be credible and rely upon reputation built on previous plays of the game. Otherwise, announcing a valiant strategy and not fulfilling it may aggravate the humiliation of the loser of the game. To commit or not to commit oneself at the beginning of the game can thus be another game of Chicken with an uncertain outcome.

Another way to deal with the result of the Chicken game is to assume that every player will choose the least-risky strategy, with the criterion called **maxi-min, in order avoid the worst possible outcome.** Since every strategy can produce a minimum value

TABLE 3.3 Chicken

The outcome of the game depends on the decision made by each boy on the basis of his expectation as to what the other boy will do.

		DEAN	
		COOPERATE C	DEFECT D
JAMES	COOPERATE C (*stop*)	Both chicken	James chicken Dean wins
	DEFECT D (*drive ahead*)	James wins Dean chicken	Both fall

(being a chicken if one cooperates, and falling into the precipice if one does not), it is assumed that the actor will choose **the strategy that guarantees the "maximum minimum" result**—in the metaphorical story: to save his life.

Note that this is also the criterion of choice that we assumed would be used by the players in the Prisoner's Dilemma game: by trying to avoid the strategy that could produce the worst possible outcome (ten years in prison), they chose to defect. In the Chicken game, in contrast, this criterion would lead the two players to cooperate—placing the outcome in the upper-left cell of Table 3.3.

In any of the possible readings of the game structure, we should expect, therefore, that there will be at least some degree of cooperation—although it is not clear who will cooperate. Note again that the worst possible outcome in which both players defect is not an equilibrium, because each player has incentives to unilaterally modify his strategy (and brake).

The Chicken game is a representation of collective action problems when unilateral cooperation is worthy because it is presumed to be able to provide *some* collective good. Logically this strategic interaction can be appropriate only for the provision of collective goods that can be provided in incomplete amounts.

Let us think about different types of collective goods from the perspectives provided by the Prisoner's Dilemma and Chicken games. In Prisoner's Dilemma–type situations, neither individual finds it beneficial to cooperate unilaterally because he would be unable to provide any amount of the collective good by himself. This assumption may fit collective goods that must be provided in entire units, such as highways, bridges, tunnels, canals, harbors, airports, schools, and hospitals. These are also called "lumpy" goods—in contrast to "continuous" goods, similar to different ways of providing sugar. To have half a bridge or a portion of a tunnel is like not having one at all. Goods of this type require some significant amount of provision—a lump—to be able to benefit their potential users, rendering partial cooperation not helpful.

In contrast, for continuous collective goods, such as certain environmental goods, unilateral cooperation can provide some amount. Partial cooperation and provision can exist. Saving the atmosphere, seas, oceans, rivers, lakes, and woods from global

BOX 3.2 BASIC GAMES

The structure of the three basic games of collective action can be represented by the following table. One actor is located in the row and the other in the column. Each actor can choose between cooperating (C) or defecting (D). The strategy of each actor leads to a collective outcome that is represented by each of the four cells in the table. If the two actors choose mutual cooperation, each receives a reward (R). If the two actors choose mutual defection, each receives a penalty (P). If one cooperates while the other defects, the cooperative actor is a sucker (S), while the defector cedes to the temptation (T).

		COLUMN	
		COOPERATE C	DEFECT D
ROW	COOPERATE C	RR	ST
	DEFECT D	TS	PP

In each game the actors prefer the outcomes in a different order. See the actors' orders of preference and the subsequent equilibrium outcomes.

GAME	ACTOR'S PREFERENCE ORDER	COLLECTIVE OUTCOMES IN EQUILIBRIUM
Prisoner's Dilemma	T > R > P > S	PP: mutual defect
Chicken	T > R > S > P	TS, ST: partial cooperation and defect
Stag Hunt	R > T > P > S	PP: mutual defect RR: mutual cooperation

warming, mentioned earlier, or the reproduction and maintenance of some animal species and plants can be done at different rates. But each of these goods can have a critical level of provision to guarantee its preservation.

Consider now the case of a river surrounded by factories polluting its water by discharging their detritus into it. The river waters can absorb rubbish from a certain number of factories and still remain usable, but when the waste from many factories surpasses some critical threshold, it results in an ecological catastrophe. Since this would be the worst result for every company making use of the river, some of them may unilaterally restrain from polluting it, even if others continue to pollute it. In this outcome some companies cooperate and others free-ride, and the subsequent benefits are suboptimal and unevenly distributed, but all are better off, including the unilateral cooperators, than if the waters ceased to be usable. Similarly, an international agreement on limiting the hunting of certain species—whales or elephants, for instance—are frequently fulfilled by some of the signatory parties to the agreement, in order to permit the reproduction of the species, while being contravened by others–in the knowledge (and righteous anger) of the former.

In some cases the real actors involved in a conflict may not be sure of whether the structure of their interaction more closely resembles the Prisoner's Dilemma or Chicken. For instance, in the previous chapter we depicted the Cold War and arms race between the United States and the Soviet Union in the second half of the twentieth century as a Prisoner's Dilemma. However, some pacifist groups in the West, fearful of the risk of massive destruction of the globe, put mutual defection at the

SOURCE 3.2 Chickens and Stags

The catchy metaphors giving ground to some basic games of collective action originated in old political speeches:

> Since the nuclear stalemate became apparent, the Governments of East and West have adopted...a policy adapted from a sport which, I am told, is practiced by some youthful degenerates. This sport is called "**Chicken**!." It is played by choosing a long straight road with a white line down the middle and starting two very fast cars towards each other from opposite ends. Each car is expected to keep the wheels of one side on the white line. As they approach each other, mutual destruction becomes more and more imminent. If one of them swerves from the white line before the other, the other, as he passes, shouts "Chicken!", and the one who has swerved becomes an object of contempt...When the game is played by eminent statesmen, they risk not only their own lives but those of many hundreds of millions of human beings.
>
> Bertrand Russell, *Common Sense and Nuclear Warfare* (1959)

> If a **stag** was to be taken, every one saw that, in order to succeed, he must abide faithfully by his post. But if a hare happened to come within the reach of any one of them, it is not to be doubted that he pursued it without scruple and, having seized his prey, worried very little if by so doing he caused his companions to lose theirs...[Over time] men may have imperceptibly acquired some rough ideas of mutual commitments and of the advantages of fulfilling them, but only to the extent that present and perceptible interest could demand it.
>
> Jean-Jacques Rousseau, *Discourse on the Origin of Inequality* (1754)

TABLE 3.4 Stag Hunt

Two hunters may either catch a hare each or cooperate to chase after a stag—a collective good to share.

		JACQUES	
		COOPERATE C	DEFECT D
JEAN	COOPERATE C *(chase stag)*	Stag to share	Jean gets nothing Jacques catches a hare
	DEFECT D *(catch hare)*	Jean catches a hare Jacques gets nothing	A hare for each

bottom of their preferences, and advocated unilateral cooperation—that is, unilateral disarmament of the United States and Western Europe even at the risk of surrendering to Soviet expansion. Along the lines of a boy in the Chicken game, afraid that his competitor would drive ahead anyway, they coined a euphonic saying: "Better red than dead." Finally, however, the West ended up neither red nor dead.

STAG HUNT

Still more likely to succeed in producing individuals' mutual cooperation are situations outlined by the game structure called Stag Hunt, or "Assurance." In this case the original story of the game refers to the attempt to organize collective action for hunting some edible animals in the woods. The basic assumption is that no individual can catch a stag by himself—that is, nobody can provide the whole collective good to be shared—but by defecting from the team, anyone can catch a hare and get some private good to eat. For the two hunters playing the game, it is also assumed that getting half a stag is better than getting a hare—that is, the individual share of the big collective item is better than what one can attain by oneself. So for each actor, it is better to cooperate and obtain the big collective good than to go out on his own, but to get something on his own is better than cooperating unilaterally and getting nothing at all (facing the stag alone).

What would you expect the outcome of the game to be? Let us call the two actors Jean and Jacques. In the fear that Jacques will go his own way to chase after a hare, Jean can do the same and obtain some good. In this outcome there would be **mutual defection**. This result can last for a long time—with lonely individuals walking through the woods each chasing after their own food. But even if this is the first, provisional outcome, **achievement of mutual cooperation should be relatively unproblematic** because all actors prefer their share of the collective good to any other outcome. Table 3.4 represents the possible strategic choices. If everybody defects, each gets a hare, but if somebody starts chasing the stag, he can expect that the other may join in the effort, and the two will obtain the collective good. The structure of this game might enlighten interactions between people with some common interest in which the participants can develop effective mutual pressure and dissuasion, leading them to cooperate. Initially, mutual defection can be an equilibrium. But some individual initiative can obtain the efficient outcome.

Quick Quiz

- Is mutual defect an equilibrium outcome in the Chicken game?

Conclusion

Game theory has helped us to analyze diverse strategic situations in which people interact to promote their common interests. From the previous discussion we can establish two basic *propositions*:

1. **Prisoner's Dilemma.** In this game, which represents the basic structure of collective action problems for the provision of public goods, each actor has a dominant strategy not to cooperate, leading to an inefficient outcome in which all the participants are worse off than if all cooperated.
2. **Tit for Tat.** The strategy of starting to cooperate and doing unto others as they do unto you can lead to sustained cooperation. Mutual cooperation is more likely the greater the uncertainty as to the length of the collective relationship and the higher the number of interactions.

Different conditions for the provision of public goods, which can be stylized with the Prisoner's Dilemma, the Chicken, or the Stag Hunt games, can entail different degrees of cooperation and defection. Nevertheless, success or failure in attaining cooperative interactions within an interest group does not necessarily correspond to socially efficient or inefficient solutions. Remember the moral ambiguity implied by the foundational story of the Prisoner's Dilemma, in which successful cooperation—among suspects of a crime—may lead them to escape punishment, while justice would most likely be implemented if they defected. The collective strength of some groups may indeed provoke conflict with other groups or favor asymmetric and biased redistributions of resources, hindering more satisfactory outcomes for greater numbers of people.

Summary

Game theory helps to analyze human interactions in strategic situations in which individual decisions depend on what each participant expects the others will do.

The Prisoner's Dilemma represents the most basic structure of the problems of collective action for the provision of public goods. Each actor has a dominant strategy not to cooperate. In the inefficient equilibrium outcome of the game, all participants are worse off than if all cooperated. This can be considered to be a representation of the free-rider problem.

The Tit for Tat meta-strategy or criterion of decision for repeated plays of the Prisoner's Dilemma is based on the idea of responding to the others' behavior. It consists of starting cooperating and doing unto others as they do unto you: cooperate when the others cooperate, and defect when the others defect. It can produce sustained cooperation when the number of interactions is indefinite.

Cooperation is more likely to hold up in certain groups the greater the uncertainty as to the length of the collective relationship and the higher the number of interactions. Iterated reciprocity may build good reputation and trust among members of the group and the building of institutions favoring further cooperation.

The Chicken game can lead to two possible outcomes, each involving some degree of cooperation among the participants. This result can be a representation of the advantages of unilateral cooperation for the provision and reproduction of "continuous" goods, such as those in question in certain environmental problems with a catastrophe threshold.

In the Stag Hunt, or "Assurance," game, mutual defection is an equilibrium outcome, but mutual cooperation is also an equilibrium outcome that should not be difficult for the participants to attain for their mutual benefit.

Key Concepts

Backward induction. Reasoning leading to mutual defection when the end of the interaction is known.

Chicken. Game with two possible outcomes in which some actors cooperate and others defect.

Competition. Action among several players each for his own benefit.

Conflict or zero-sum game. An intercation in which the gains for some people imply losses for others.

Cooperation. Action for the common benefit.

Coordination game. An interaction in which there are efficient results for all participants.

Dominant strategy. The best actor's decision regardless of the other actor's decision.

Equilibrium. Stable outcome of a game. In Nash's concept, an outcome from which no actor has incentives to move away by changing his strategy unilaterally.

Maxi-min. Criterion of decision to prevent the worst result and guarantee the "maximum minimum."

Positive or non-zero-sum game. An interaction in which mutual cooperation can produce gains for all participants.

Prisoner's Dilemma. Non-zero-sum game with an inefficient equilibrium outcome and in which nobody cooperates.

Stag Hunt. An assurance game with mutual defection and mutual cooperation equilibrium outcomes.

Tit for Tat. Meta-strategy for repeated plays of the Prisoner's Dilemma that can produce sustained cooperation.

Questions for Review

1. What's the equilibrium outcome of the Prisoner's Dilemma game?
2. What's "Tit for Tat?"
3. What's the outcome of the Chicken game?
4. What's the outcome of the Stag Hunt game?

Problems and Applications

1. Play this coordination game: After being separated during a term break, you are to meet a fellow student on campus on the first day of classes at twelve noon. You did not make an arrangement regarding where to meet and your cell phone has been stolen. You both have to guess where to meet and try to make your guesses coincide. Note that the two of you have identical interests, but you will succeed only if you make identical choices. You win only if you do what the other expects you to do.

   Collect the written responses and match each student's response with those of the others.
2. Represent the following strategic situation with the help of game theory: At a beach resort in California is the Hotel Flamingo which produces waste and refuse polluting the waters of the sea. Usually Mr. Grant, a hotel employee, quite successfully does the job of filtering the waste. Since about a year ago, a new tourist resort, PalmHouse, which opened just a few yards away, has been producing similar waste, polluting the waters as well. Initially, Mr. Grant did his best to filter everything he could. The result was not optimal, however, since it would have been better for PalmHouse to have organized its own filtering mechanisms. Some visitors noticed dirty waters at the beach, but there was no visible decrease in the number of customers. However, Mr. Grant retired a few weeks ago, and now nobody does his job. There is much concern regarding the visible water pollution. The tourist season

seems to be in danger. Everybody thinks that somebody should do something about it.

a. Define the public good as provided by actors' cooperation.
b. Order actors' preferences regarding all possible outcomes.
c. Draw a table representing the possible outcomes.
d. Discuss the actors' strategies and identify the outcome of the game. What may it mean in terms of the provision (or not) of the public good?

3. Represent the following strategic situation with the help of game theory: One individual is assaulting female residents in suburban Town. The police are not acting appropriately. The victims think that just two men would be sufficient to apprehend the criminal. A voluntary night patrol is organized by the president of the residents' association. Three men from Town will be called each night, at random. About eight thousand families live in the area. The president sends a letter to each family head explaining the plan and allocating one patrol night to each addressee. On the first day, three family heads are called. Each of them has to decide whether to attend.
   a. Define the public good as provided by actors' cooperation.
   b. Order actors' preferences regarding all possible outcomes.
   c. Draw a table representing all possible outcomes.
   d. Discuss actors' strategies and identify the outcome of the game. What does it mean in terms of provision or not of the public good?
4. Play the repeated Prisoner's Dilemma game and register your scores by using any of these websites:
   www.gametheory.net/Web/PDilemma
   www.people.bath.ac.uk/mk213/ipd
   www.paulspages.co.uk/hmd
5. Working in pairs, play the repeated Prisoner's Dilemma. Two of you can sit at the front of the classroom, backs to each other, and at successive indications from the instructor, you can raise either a green card, indicating cooperation (C), or a red card, indicating defection (D). Both players should raise their cards simultaneously, without being able to see the other's action in anticipation. The game must be played multiple times until the instructor ends it. The scores are the following:
   - One player chooses D (5 points) while the other chooses C (0 points).
   - Both choose C (3 points each).
   - Both choose D (1 point each).
   - First, make the following decisions:
   - Start with either C or D.
   - Retaliate (or not) to the other player's decision at the next play.
   - Remember past grievances or forget and forgive. Remember that to win the match you need not win every play.
   - Show a clear criterion of conduct or, rather, be confusing and unpredictable.

   At the beginning of the game you can announce what you are doing to do (or what you intend the other player to believe you are going to do). Remain silent throughout the rest of the game, taking into account only the other player's decisions.

   Somebody should record the scores of each play and make the final sum at the end of the game. Observe whether mutual cooperation emerges and becomes more frequent with the play of the game. The results may be different if the players know how many times they are going to play their interaction or not.