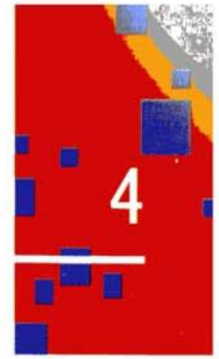


EVOLUTIONARY GAMES



EVOLUTIONARY GAME THEORY means that the fitness of individuals is not constant, but depends on the relative proportions (frequencies) of the different phenotypes in the population: fitness is frequency dependent. Evolutionary game theory is the generic approach to evolutionary dynamics and contains as a special case constant selection.

Game theory was invented by John von Neumann and Oskar Morgenstern. They wanted to design a mathematical theory to study human behavior in strategic and economic decisions. Von Neumann was a Hungarian-born mathematician working at the Institute for Advanced Study, where he invented and revolutionized several fields of mathematics. We have already encountered him in Chapter 3 in connection with the terms “translation” and “transcription,” which he invented when thinking about how to conceive of a machine that could reproduce itself. He built the first computer that held the program for the calculation in its memory rather than in its hardware. Incidentally, one of the first projects that this computer did in its spare time was a mathematical simulation of an evolutionary system.

John Nash came as a mathematics Ph.D. student to Princeton University with a remarkably short letter of reference: “This man is a genius.” Nash invented a simple but important concept in game theory, which is now called the “Nash equilibrium.” A Nash equilibrium is very similar to an evolutionarily stable strategy (ESS). Both concepts are important for evolutionary dynamics. Nash’s Ph.D. thesis led to a single-page paper in the *Proceedings of the National Academy of Sciences USA* (1950), which earned him a Nobel Prize in economics in 1994.

William Hamilton and Robert Trivers were among the first to use game theoretical ideas in biology, but the field of evolutionary game theory was founded by the work of John Maynard Smith and others, including Peter Taylor, Josef Hofbauer, and Karl Sigmund.

John Maynard Smith and George Price introduced game theory to evolutionary biology and population thinking to game theory in a paper published in *Nature* in 1973. Traditional game theory typically analyzes an interaction between two players, for example, you and me. The question is how you can maximize your payoff in a game, given that you do not know what I will do. The concept of rationality comes into play. You may assume that I will act in order to maximize my payoff. Given this assumption, you will then behave to maximize your payoff. But nothing guarantees that I will behave rationally, and in fact many experimental games show that humans do not behave rationally.

Evolutionary game theory does not rely on rationality. Instead it considers a population of players interacting in a game. Individuals have fixed strategies. They interact randomly with other individuals. The payoffs of all these encounters are added up. Payoff is interpreted as fitness, and success in the game is translated into reproductive success. Strategies that do well reproduce faster. Strategies that do poorly are outcompeted. This is straightforward natural selection.

In Figure 4.1 we see two phenotypes. *A* can move while *B* cannot. *A* pays a certain cost for the ability to move, but also gains the associated advantage. Suppose the cost-benefit analysis leads to a fitness of 1.1 for *A* compared to a fitness of 1 for *B*. In this setting, fitness is constant, and *A* will certainly outcompete *B*. But imagine that the advantage of being able to move is large when few others are on the road, but diminishes as the highways get blocked up. In this case, the fitness of *A* is not constant, but is a declining function

Constant selection:



Frequency-dependent selection:

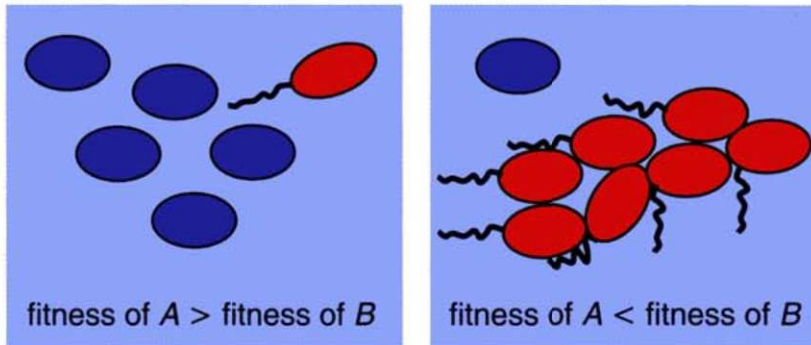


Figure 4.1 Constant selection means fitness neither depends on the composition of the population nor changes over time. For example, A has constant fitness 1.1, while B has constant fitness 1. In contrast, frequency-dependent selection means that fitness does depend on the relative abundance (= frequency) of individual types. A has the ability to move. If few other cells are moving, then A has a larger fitness than B . But if many other cells “are on the road,” this fitness advantage is reversed (in this hypothetical example).

of the frequency of A . A has a higher fitness than B when A is rare, but has a lower fitness than B when A is common. What is the outcome of such a selection process?

Let us formalize the general case of frequency-dependent selection between two strategies A and B . Denote by x_A the frequency of A and by x_B the frequency of B . The vector $\vec{x} = (x_A, x_B)$ defines the composition of the population. Denote by $f_A(\vec{x})$ the fitness of A and by $f_B(\vec{x})$ the fitness of B . The selection dynamics can be written as

$$\begin{aligned} \dot{x}_A &= x_A[f_A(\vec{x}) - \phi] \\ \dot{x}_B &= x_B[f_B(\vec{x}) - \phi] \end{aligned} \tag{4.1}$$

The average fitness is given by $\phi = x_A f_A(\vec{x}) + x_B f_B(\vec{x})$.

Because $x_A + x_B = 1$ at all times, we can introduce the variable x with $x_A = x$ and $x_B = 1 - x$. We can write the fitness functions as $f_A(x)$ and $f_B(x)$.

Frequency-dependent selection of two strategies, A and B

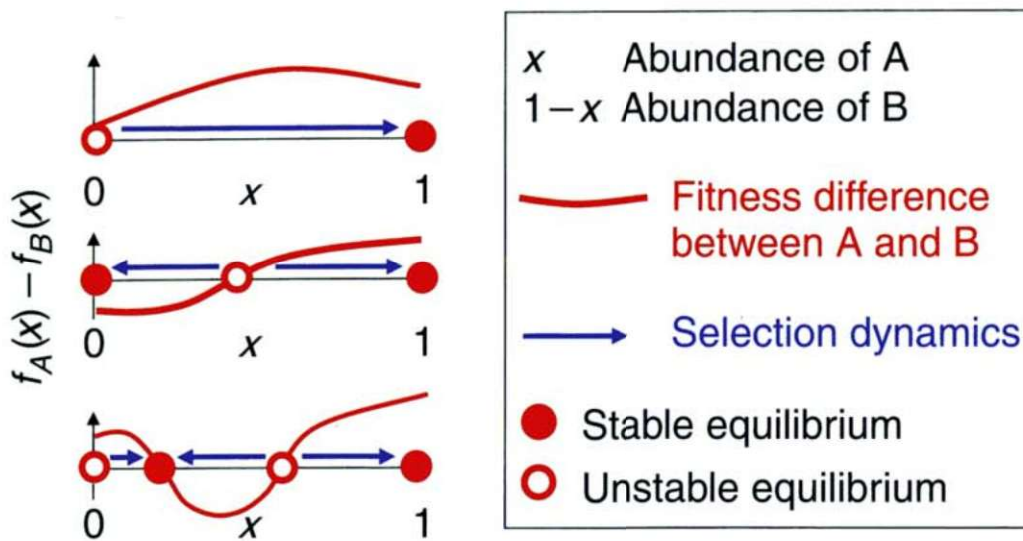


Figure 4.2 Frequency-dependent selection between two strategies can lead to multiple stable and unstable equilibria. The red curve, $f_A(x) - f_B(x)$, indicates the fitness difference between A and B as a function of the frequency, x , of A . If $f_A(x) - f_B(x) > 0$, then the frequency of A will increase, as shown by the blue arrows indicating the direction of selection dynamics. If $f_A(x) - f_B(x) < 0$, then the frequency of A will decline. Whenever $f_A(x) - f_B(x) = 0$, the frequency of A will not change. This condition characterizes an equilibrium of selection dynamics. If the slope of $f_A(x) - f_B(x)$ is positive at this point, then the equilibrium is unstable. If the slope is negative, then the equilibrium is stable. The points $x = 0$ and $x = 1$ are always equilibria. The equilibrium $x = 0$ is stable if $f_A(0) - f_B(0) < 0$. The equilibrium $x = 1$ is stable if $f_A(1) - f_B(1) > 0$.

System (4.1) leads to

$$\dot{x} = x(1-x)[f_A(x) - f_B(x)]. \quad (4.2)$$

The equilibria of this differential equation are given by $x = 0$, $x = 1$, and all values $x \in (0, 1)$ that satisfy $f_A(x) = f_B(x)$. The equilibrium $x = 0$ is stable if $f_A(0) < f_B(0)$. Conversely, the equilibrium $x = 1$ is stable if $f_A(1) > f_B(1)$. An interior equilibrium, x^* , is stable if the derivatives of the functions f_A and f_B satisfy $f'_A(x^*) < f'_B(x^*)$. Figure 4.2 gives a graphical representation. There can be several stable and unstable equilibria in the interior of the interval $[0, 1]$.

Payoff matrix

	A	B	
A	<i>a</i>	<i>b</i>	If A meets A, both get <i>a</i> .
B	<i>c</i>	<i>d</i>	If B meets B, both get <i>d</i> . If A meets B, A gets <i>b</i> and B gets <i>c</i> .

Figure 4.3 A game between two strategies, A and B, is defined by a 2×2 payoff matrix.

4.1 TWO-PLAYER GAMES

Normally, a game with two strategies, A and B, is described by a payoff matrix

$$\begin{matrix} & A & B \\ A & \begin{pmatrix} a & b \end{pmatrix} \\ B & \begin{pmatrix} c & d \end{pmatrix} \end{matrix} \quad (4.3)$$

The payoff matrix is read in the following way: A gets payoff a when playing against A; A gets payoff b when playing against B; B gets payoff c when playing against A; B gets payoff d when playing against B (Figure 4.3).

The key idea of evolutionary game theory is to consider a population of A and B players and to equate payoff with fitness. If x_A is the frequency of A and x_B the frequency of B, then the expected payoff for A and B is respectively given by

$$\begin{aligned} f_A &= ax_A + bx_B \\ f_B &= cx_A + dx_B \end{aligned} \quad (4.4)$$

These equations assume that for each player the probability of interacting with an A player is x_A and the probability of interacting with a B player is x_B . Thus players meet each other randomly.

Frequency-dependent selection dynamics between two strategies, A and B .

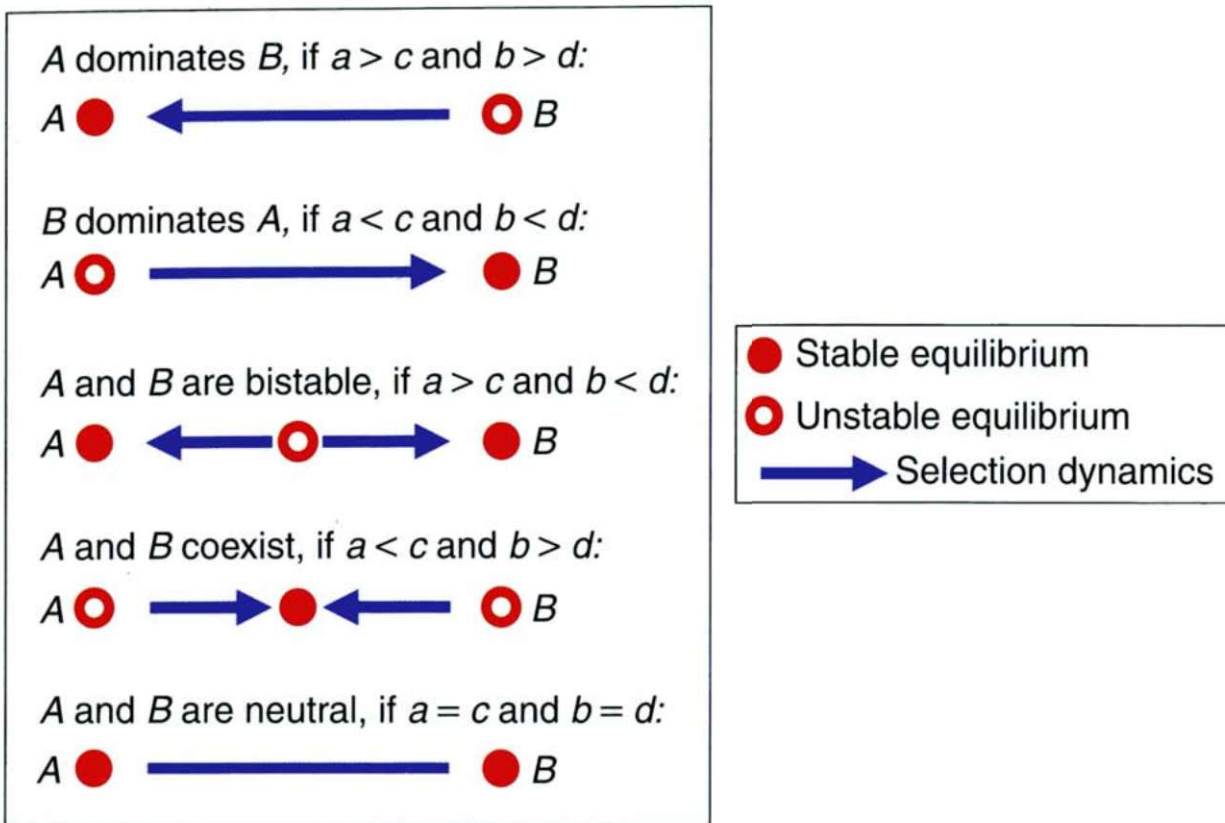


Figure 4.4 There are five possibilities for the selection dynamics between two strategies: (i) A dominates B , (ii) B dominates A , (iii) A and B are bistable, (iv) A and B coexist in a stable equilibrium, and (v) A and B are neutral variants of each other.

Let us now insert these linear fitness functions into equation (4.1). Let $x = x_A$ as before. We obtain

$$\dot{x} = x(1-x)[(a-b-c+d)x + b-d]. \quad (4.5)$$

We want to classify the behavior of the nonlinear differential equation (4.5) depending on the ranking of the entries in the payoff matrix (4.1). There are five cases (Figure 4.4):

(i) A dominates B . This is the case if $a > c$ and $b > d$. If you were to play this game with another person, then no matter whether the other person chooses A or B it is best for you to play A . For a population of A and B players, this ranking of payoff values implies that the average fitness of A will always

exceed that of B . Thus selection will favor A over B for any composition of the population. Selection will lead to the state where the whole population consists of A players, $x_A = 1$. More accurately, we say that A dominates B if $a \geq c$ and $b \geq d$, where at least one inequality must be strict.

(ii) B dominates A . This is the case if $a < c$ and $b < d$. This situation is the mirror image of case (i) with A and B exchanged. Again, more accurately, we say that A dominates B if $a \leq c$ and $b \leq d$, where at least one inequality must be strict.

(iii) A and B are bistable. This is the case if $a > c$ and $b < d$. If you were to play this game with another person, you should try to play the same choice as the other person. A is the best response for A . B is the best response for B . For the selection dynamics within a population, the outcome depends on the initial condition. There is an unstable equilibrium in the interior of the interval $[0, 1]$ given by $x^* = (d - b)/(a - b - c + d)$. If the initial condition, $x(0)$, is less than this value, $x(0) < x^*$, then the system will converge to all- B . If $x(0) > x^*$ then the system will converge to all- A .

(iv) A and B stably coexist. This is the case if $a < c$ and $b > d$. If you were to play this game with another person, you should always try to choose the opposite of what your opponent does. A is best response to B , and B is best response to A . A population of A and B players will converge to the interior, stable equilibrium

$$x^* = \frac{d - b}{a - b - c + d}. \quad (4.6)$$

(v) A and B are neutral. This is the case if $a = c$ and $b = d$. If you were to play this game with another person, then no matter what you chose you would always have exactly the same payoff as your opponent. Selection will not change the composition of the population. Any mixture of A and B is an equilibrium for selection dynamics.

4.2 THE NASH EQUILIBRIUM

The Nobel Prize-winning concept of a Nash equilibrium is defined in the following way. Imagine there is a game between two people. If both play a strategy that happens to be a Nash equilibrium, then neither person can deviate from this strategy and increase his payoff.

Consider the general payoff matrix between two strategies, A and B ,

$$\begin{matrix} & A & B \\ A & \begin{pmatrix} a & b \end{pmatrix} \\ B & \begin{pmatrix} c & d \end{pmatrix} \end{matrix}$$

We have the following criteria:

- (i) A is a strict Nash equilibrium if $a > c$.
- (ii) A is a Nash equilibrium if $a \geq c$.
- (iii) B is a strict Nash equilibrium if $d > b$.
- (iv) B is a Nash equilibrium if $d \geq b$.

Let us explore the following game

$$\begin{matrix} & A & B \\ A & \begin{pmatrix} 3 & 0 \end{pmatrix} \\ B & \begin{pmatrix} 5 & 1 \end{pmatrix} \end{matrix} \tag{4.7}$$

If both players choose A , then one player can improve his payoff by switching to B . If both play B , then neither player can improve his payoff by switching to A . Hence B is a Nash equilibrium. Note that A is dominated by B . Note also that playing the Nash equilibrium, B , in this game leads to a lower payoff than playing the dominated strategy, A . The payoff matrix (4.7) is an example of the famous Prisoner's Dilemma game, which we will study in Chapter 5.

Consider the game

$$\begin{matrix} & A & B \\ A & \begin{pmatrix} 3 & 1 \end{pmatrix} \\ B & \begin{pmatrix} 5 & 0 \end{pmatrix} \end{matrix} \tag{4.8}$$

If both players choose A , then one player can improve his payoff by switching to B . If both play B , then again one player can improve his payoff by switching

to A . Hence neither A nor B is a Nash equilibrium. This is an example of a Hawk-Dove game, which will be studied in section 4.6.

Finally, consider the game

$$\begin{array}{c} \\ A \\ B \end{array} \begin{array}{cc} A & B \\ \left(\begin{array}{cc} 5 & 0 \\ 3 & 1 \end{array} \right) \end{array} \quad (4.9)$$

If both players choose A , then neither player can improve his payoff by switching to B . If both play B , then again neither player can improve his payoff by switching to A . Hence both A and B are Nash equilibria.

4.3 EVOLUTIONARILY STABLE STRATEGY (ESS)

John Maynard Smith invented the important concept of an evolutionarily stable strategy when he was unaware of the Nash equilibrium in game theory. Imagine a large population of A players. A single mutant of type B is introduced. The game between A and B is given by the general payoff matrix (4.3), and the fitness functions are given by (4.4). What is the condition for selection to oppose the invasion of B into A ?

Let us assume there is an infinitesimally small quantity of B invaders. Thus the frequency of B is ϵ . The frequency of A is $1 - \epsilon$. For this population, the fitness of A is greater than the fitness of B if

$$a(1 - \epsilon) + b\epsilon > c(1 - \epsilon) + d\epsilon. \quad (4.10)$$

Canceling the ϵ terms, this inequality leads to

$$a > c. \quad (4.11)$$

If, however, it happens that $a = c$, then inequality (4.10) leads to

$$b > d. \quad (4.12)$$

Therefore we summarize: strategy A is ESS if either (i) $a > c$ or (ii) $a = c$ and $b > d$. This definition guarantees that selection will oppose the invasion

of B into A . As we will see later, this concept holds only for infinitely large populations and for infinitesimally small quantities of the potential invader.

4.4 MORE THAN TWO STRATEGIES

Let us explore games with more than two strategies. The payoff for strategy S_i versus S_j is given by $E(S_i, S_j)$.

(i) Strategy S_k is a strict Nash equilibrium if

$$E(S_k, S_k) > E(S_i, S_k) \quad \forall i \neq k \quad (4.13)$$

The symbol \forall means “for all.” Thus $\forall i \neq k$ reads “for all i not equal to k .”

(ii) Strategy S_k is a Nash equilibrium if

$$E(S_k, S_k) \geq E(S_i, S_k) \quad \forall i \quad (4.14)$$

(iii) Strategy S_k is ESS, if $\forall i \neq k$ we have either

$$E(S_k, S_k) > E(S_i, S_k) \quad (4.15)$$

or

$$E(S_k, S_k) = E(S_i, S_k) \quad \text{and} \quad E(S_k, S_i) > E(S_i, S_i). \quad (4.16)$$

Note that ESS guarantees that selection will oppose any potential invader. The same is true for a strict Nash equilibrium, but not for a Nash equilibrium. If $E(S_k, S_k) = E(S_j, S_k)$ and $E(S_k, S_j) < E(S_j, S_j)$ then S_k is still a Nash equilibrium, but selection will favor S_j invading S_k . Thus it makes sense to add a fourth definition.

(iv) Strategy S_k is stable against invasion by selection (let us call this “weak ESS”) if $\forall i \neq k$ we have either

$$E(S_k, S_k) > E(S_i, S_k) \quad (4.17)$$

or

$$E(S_k, S_k) = E(S_i, S_k) \quad \text{and} \quad E(S_k, S_i) \geq E(S_i, S_i). \quad (4.18)$$

If a strategy is a strict Nash equilibrium then it is also an ESS. If a strategy is an ESS then it is also a weak ESS. If a strategy is a weak ESS then it is also a Nash equilibrium. Thus strict Nash implies ESS implies weak ESS implies Nash:

$$\text{strict Nash} \Rightarrow \text{ESS} \Rightarrow \text{weak ESS} \Rightarrow \text{Nash}. \quad (4.19)$$

All of these concepts play important roles when studying the evolutionary dynamics of frequency-dependent selection.

The concept of an “unbeatable strategy” was introduced by William Hamilton in his work on sex ratios preceding John Maynard Smith’s work on evolutionary game theory. Strategy S_k is unbeatable if $\forall i \neq k$:

$$E(S_k, S_k) > E(S_i, S_k) \quad \text{and} \quad E(S_k, S_i) > E(S_i, S_i). \quad (4.20)$$

Therefore an unbeatable strategy dominates every other strategy. An unbeatable strategy is certainly a strict Nash equilibrium. An unbeatable strategy is the most you can ask for, but usually you are asking for too much. Unbeatable strategies are rare.

4.5 REPLICATOR DYNAMICS

Peter Taylor and Leo Jonker were the first to introduce a differential equation for evolutionary game dynamics. They were quickly followed by Christopher Zeeman (Warwick) as well as Peter Schuster, Josef Hofbauer, and Karl Sigmund (all three in Vienna).

Given what we have seen so far in this book, the equation that we will write down now is an obvious next step. Consider the interaction among n strategies. The payoff for strategy i when interacting with strategy j is given by a_{ij} . The $n \times n$ -matrix $A = [a_{ij}]$ is called the “payoff matrix.” Let x_i denote the frequency of strategy i . The expected payoff of strategy i is given by $f_i = \sum_{j=1}^n x_j a_{ij}$. The average payoff is given by $\phi = \sum_{i=1}^n x_i f_i$. Equating payoff with fitness, we obtain the replicator equation (Figure 4.5)

$$\dot{x}_i = x_i(f_i - \phi) \quad i = 1, \dots, n \quad (4.21)$$

The replicator equation

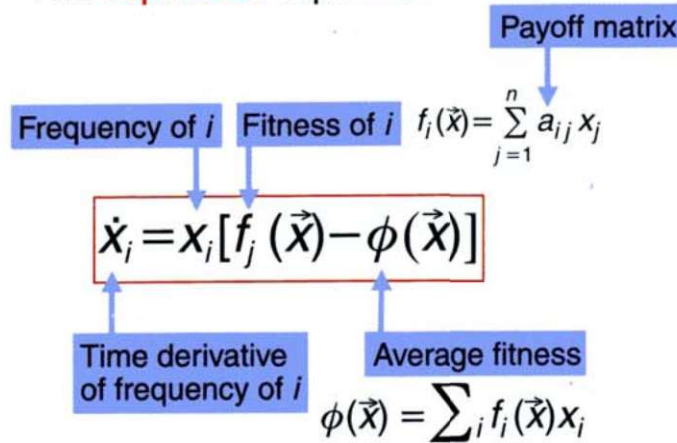


Figure 4.5 The replicator equation is the cornerstone of evolutionary game dynamics. It describes frequency-dependent selection among n different phenotypes (strategies) in an infinitely large population without mutation. In general, the fitness, f_i , of the phenotype, i , depends on the composition of the population, $\vec{x} = (x_1, \dots, x_n)$. Usually f_i is a linear function of the frequencies x_j , and the coefficients of this function are the entries of the payoff matrix $A = [a_{ij}]$. The entry a_{ij} denotes the payoff for strategy i interacting with strategy j .

The difference between (4.21) and (2.16) is frequency-dependent as opposed to constant selection. The fitness values are linear functions of the frequencies.

Equation (4.21) is defined on the simplex S_n , which is given by $\sum_{i=1}^n x_i = 1$. We note that the interior of the simplex is invariant: if a trajectory starts in the interior, it will always remain in the interior; it might converge to the boundary, but it will never actually reach the boundary. Moreover, we note that each face of the simplex is invariant. A face is a subset of the simplex where at least one strategy has zero frequency. A strategy which is not there will not appear.

Replicator dynamics describe pure selection without mutation. Often, however, we ask the question whether a new strategy, which is not present in the population, could invade the population and increase in frequency. Thus, while not explicitly modeled, mutation is on the mind of the analyst.

The corners (vertices) of the simplex are fixed points of the replicator dynamics. Depending on the payoff matrix, A , there can be fixed points in the interior and in every face of the simplex.

4.5.1 Two Strategies

For $n = 2$, we have already derived a complete classification of all possible evolutionary dynamics in section 4.2. The simplex S_2 is the closed interval $[0, 1]$. The corner points, $x_1 = 0$ and $x_1 = 1$, are always fixed points of the replicator dynamics. In the interior, there can be at most one isolated fixed point. Such a fixed point exists if $(a_{11} - a_{21})(a_{12} - a_{22}) < 0$. This condition ensures that neither A dominates B nor B dominates A . Instead A and B can either be bistable or coexist. The interior fixed point is stable if A and B coexist which is the case for $a_{11} < a_{21}$ and $a_{12} > a_{22}$.

The neutral case, $a_{11} = a_{21}$ and $a_{12} = a_{22}$, implies that every point of the interval $[0, 1]$ is an equilibrium point. In this case, the replicator dynamics do not move. The strategies A and B have identical fitness for any composition of the population. In such a case, we would say that replicator dynamics do not adequately describe the biological behavior. For any finite population size, the relative proportions of A and B will drift until eventually one strategy becomes extinct. We will discuss the evolutionary dynamics of finite populations in Chapter 6.

Constant selection between A and B is obtained as a special case of the replicator equation if $a_{11} = a_{12} \neq a_{21} = a_{22}$. Thus evolutionary game dynamics is the generic description of natural selection, with constant selection being a special case.

4.5.2 Three Strategies

For $n = 3$, interesting new dynamical features are possible. The phase space is the simplex S_3 , which can be represented by a triangle with identical sides. There is a complete classification of all possible phase portraits.

Of special interest is the case where A is dominated by B , B is dominated by C , and C is dominated by A . This is the so-called Rock-Paper-Scissors game. In this well-known children's game, rock beats scissors which beat paper which beats rock.

Any 3×3 payoff matrix with cyclic domination of the three strategies characterizes a Rock-Paper-Scissors game. The analysis of the evolutionary game dynamics can be simplified by a trick that is applicable to any replicator equation: the dynamics of the replicator equation (4.21) remain unchanged if an

arbitrary constant is added to each entry in a column of the payoff matrix. Thus by subtracting the diagonal element from each column, we can transform every payoff matrix into a payoff matrix that has only zero entries in the diagonal.

For example, the payoff matrix

$$A = \begin{matrix} & \begin{matrix} R & S & P \end{matrix} \\ \begin{matrix} R \\ S \\ P \end{matrix} & \begin{pmatrix} 4 & 2 & 1 \\ 3 & 1 & 3 \\ 5 & 0 & 2 \end{pmatrix} \end{matrix}$$

can be transformed to

$$A = \begin{matrix} & \begin{matrix} R & S & P \end{matrix} \\ \begin{matrix} R \\ S \\ P \end{matrix} & \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix} \end{matrix} \quad (4.22)$$

Both payoff matrices lead to identical replicator dynamics.

The payoff matrix (4.22) defines the symmetric Rock-Paper-Scissors game. The interior of the simplex S_3 contains a unique equilibrium point given by $(1/3, 1/3, 1/3)$. This point is stable, but not asymptotically stable. There are infinitely many periodic orbits surrounding this center. In fact all other points in the interior are on periodic orbits. The time average of each cycle is given by $(1/3, 1/3, 1/3)$. This situation is not generic: small deviations from the symmetry of A will change the phase portrait.

Note that payoff matrix (4.22) describes a zero-sum game. For a zero-sum game we have $a_{ij} = -a_{ji}$. The gain of one player is the loss of another; the average fitness of the population is always zero, $\phi = 0$. Thus the replicator equation becomes $\dot{x}_i = x_i f_i$ for all $i = 1, \dots, n$.

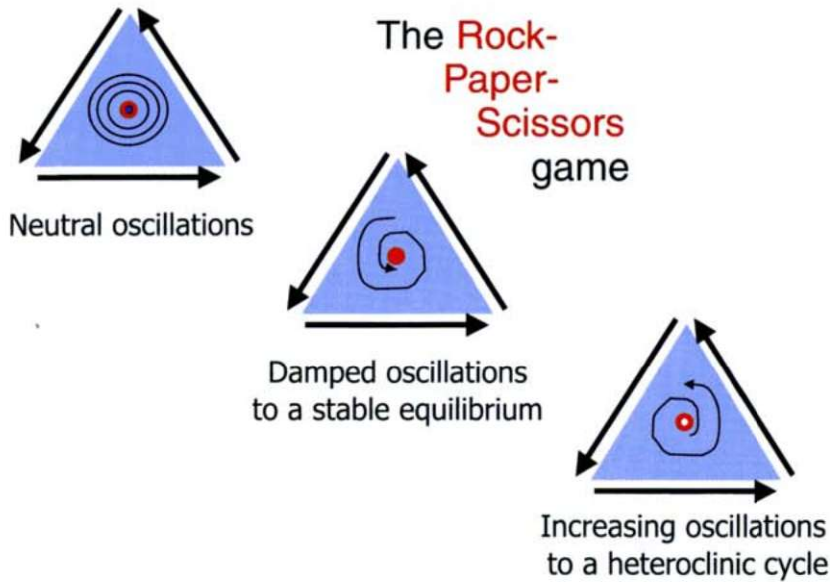


Figure 4.6 In the Rock-Paper-Scissors game, there is cyclic domination among three strategies. There is always one interior equilibrium. Depending on the payoff matrix, this interior equilibrium is (i) a center surrounded by neutral oscillations, (ii) a stable equilibrium that is approached in damped oscillations, or (iii) an unstable equilibrium, in which case there are oscillations of increasing amplitude converging to the boundary of the simplex.

The general Rock-Paper-Scissors game is given by the payoff matrix

$$A = \begin{pmatrix} 0 & -a_2 & b_3 \\ b_1 & 0 & -a_3 \\ -a_1 & b_2 & 0 \end{pmatrix} \quad (4.23)$$

There are two possibilities.

(i) if the determinant of A is positive (which is the case for $a_1 a_2 a_3 < b_1 b_2 b_3$), then there exists a unique interior equilibrium that is globally stable. The trajectories of the replicator dynamics starting from any interior initial condition will converge to this equilibrium in an oscillatory manner (Figure 4.6).

(ii) if the determinant of A is negative (which is the case for $a_1 a_2 a_3 > b_1 b_2 b_3$), then there exists a unique interior equilibrium that is unstable. The trajectories of the replicator dynamics starting from any interior initial condition will converge to the boundary of the simplex in oscillations with increasing amplitude. The boundary of the simplex is a heteroclinic cycle that is an attractor for all trajectories starting in the interior (a heteroclinic cycle is an invariant set containing saddle points on trajectories connecting them). For

the differential equation, the oscillations will converge to the heteroclinic cycle without ever reaching it. For practical purposes, one of the three strategies will eventually become extinct, either by a rounding error of the computer program or by a random fluctuation in nature. Then only two strategies are left, and one will outcompete the other, leaving a single strategy in the end. It cannot be predicted which strategy will win. Thus the attracting heteroclinic cycle reveals a kind of deterministic unpredictability that is of a different type than chaos.

4.5.3 More than Three Strategies

For four species, the replicator dynamics are defined on the simplex S_4 given by a tetrahedron. This is the minimum number of dimensions that are needed for the replicator equation to allow limit cycles and chaotic attractors.

Finally, let us state some important results that hold in any dimension.

(i) An interior equilibrium of the replicator dynamics is given by the solution of the linear system of equations

$$f_1 = f_2 = \cdots = f_n \quad \text{and} \quad x_1 + x_2 + \cdots + x_n = 1. \quad (4.24)$$

Remember that $f_i = \sum_{j=1}^n a_{ij}x_j$. This system of n linear equations has either one or zero nondegenerate solutions. Thus there can be at most one isolated equilibrium in the interior of the simplex S_n .

(ii) If there is no equilibrium in the interior, then all trajectories converge to the boundary of the simplex. Thus there can be no chaotic attractor and no limit cycle in the interior if there is no equilibrium in the interior. This result is very helpful, because sometimes it is possible to show that a particular replicator equation admits no interior equilibrium. If this is the case, then we know that nothing more complicated can happen in the interior; coexistence of all strategies is impossible.

(iii) In degenerate cases, the replicator equation can admit a manifold of equilibria in the interior. These equilibria can be stable, but not asymptotically stable.

(iv) If a strategy is a strict Nash equilibrium or ESS, then the corner point of the simplex, which corresponds to a pure population of this strategy, is asymptotically stable.

Proofs of these assertions are given in the book by Hofbauer and Sigmund (1998).

4.6 HAWK OR DOVE?

Animals of the same species fight with each other. Conflicts arise over food, territory, or sex. Female lions fight to maintain their hunting grounds. Male chimpanzees fight for dominance of a group: the alpha male has to withstand challenges from other males, and in return gets the majority of matings. There are many examples of fierce and even deadly conflicts among animals. The human species, of course, is unsurpassed in its sad efficiency in inflicting murder, war, and genocide.

Often, however, conflicts between animals (including humans) do not escalate, but are fought out with those involved obeying certain limits. Ethology has a long-standing fascination with so-called conventional fights. A sequence of threatening signals and displays allow the contestants to assess each other's strength or determination before one of them simply walks away. Stags perform roaring matches followed by parallel walks and head pushing with interlocked antlers. Despite the lethal points of their antlers, only few contests lead to serious injury.

For a long time, biologists have accepted the following explanation: conventional fighting is frequently observed in nature because it is good for the species. Fighting that leads to serious injury is bad for the species. This argument, however, is problematic. There can certainly be selection between groups or whole species, but often stronger selection is put on the shoulders of individuals. If one individual in a population were to disobey the rules and escalate all fights by inflicting serious injury, then it might win many contests and thereby reproduce its genes more efficiently than others would.

Conventional fighting from the perspective of individual selection was first analyzed by John Maynard Smith. There are two strategies, hawks (H) and

doves (D). While hawks escalate fights, doves retreat when the opponent escalates. The benefit of winning the fight is b . The cost of injury is c . If two hawks meet, then the expected payoff for each of them is $(b - c)/2$. The fight will escalate. One hawk wins, while the other is injured. Since both hawks are equally strong, the probability of winning or losing is $1/2$. If a hawk meets a dove, the hawk wins and receives payoff b , while the dove retreats and receives payoff 0. If two doves meet, there will be no injury. One of them will eventually win. The expected payoff is $b/2$. Thus, the payoff matrix is given by

$$\begin{array}{c} \text{H} \quad \text{D} \\ \text{H} \left(\begin{array}{cc} \frac{b-c}{2} & b \\ 0 & \frac{b}{2} \end{array} \right) \\ \text{D} \end{array} \quad (4.25)$$

If $b < c$, then neither strategy is a Nash equilibrium. If everybody else plays “hawk,” it is best to play “dove.” If everybody else plays “dove,” it is best to play “hawk.” Thus hawks and doves can coexist. Selection dynamics will lead to a mixed population. At the stable equilibrium, the frequency of hawks is given by b/c . If the cost of injury is much larger than the benefit of winning the fight, $c \gg b$, then the equilibrium frequency of hawks will be small.

The name of the game is somewhat misleading, because normally we think of games between individuals of the same species. Moreover, real doves fight to death when confined in a cage.

4.6.1 Mixed Strategies

So far we have considered pure strategies that either play hawk all the time or play dove all the time. Let us now consider mixed strategies which play hawk with probability p and dove with probability $1 - p$. There is no longer a discrete set of just two strategies but a continuous set of infinitely many strategies. The space of strategies is given by the closed interval $[0, 1]$. The payoff for strategy p_1 versus p_2 is given by

$$E(p_1, p_2) = \frac{b}{2}(1 + p_1 - p_2 - \frac{c}{b}p_1p_2). \quad (4.26)$$

From this function we can infer that the strategy $p^* = b/c$ is evolutionarily stable. Note that

$$\begin{aligned}
 E(p^*, p^*) &= \frac{b}{2} \left(1 - \frac{b}{c}\right) \\
 E(p, p^*) &= \frac{b}{2} \left(1 - \frac{b}{c}\right) \\
 E(p^*, p) &= \frac{b}{2} \left(1 + \frac{b}{c} - 2p\right) \\
 E(p, p) &= \frac{b}{2} \left(1 - \frac{c}{b} p^2\right)
 \end{aligned}
 \tag{4.27}$$

We see that $E(p^*, p^*) = E(p, p^*)$ for all p . Therefore p^* is a Nash equilibrium, but not a strict Nash equilibrium. Since $E(p^*, p) > E(p, p)$ for all $p \neq p^*$, it follows that p^* is an evolutionarily stable strategy.

4.7 THERE IS ALWAYS A NASH EQUILIBRIUM

Consider a game given by an $n \times n$ payoff matrix A . There are n pure strategies, S_1, \dots, S_n . Depending on the details of A , there may or may not exist a Nash equilibrium among those pure strategies. There is, however, always a Nash equilibrium if one considers the set of all mixed and all pure strategies. This is a deep result in game theory.

A strategy is given by the vector $\vec{p} = (p_1, \dots, p_n)$. Here p_i is the probability of playing S_i . Clearly $\sum_{i=1}^n p_i = 1$. The payoff for strategy \vec{q} versus strategy \vec{p} is given by

$$E(\vec{q}, \vec{p}) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} q_i p_j.
 \tag{4.28}$$

In vector notation we can write

$$E(\vec{q}, \vec{p}) = \vec{q} A \vec{p}.
 \tag{4.29}$$

It can be shown that for any payoff matrix A there exists at least one strategy \vec{q} with the property that

$$\vec{q}A\vec{q} \geq \vec{p}A\vec{q} \quad \forall \vec{p} \quad (4.30)$$

Thus \vec{q} is a best reply to itself, a Nash equilibrium.

4.8 CHICKEN AND SNOWDRIFT

In the Chicken game, two cars head for each other at high speed. The loser is whoever chickens out first. The winner stays on the track. If neither driver chickens out, there is the substantial cost of a collision. Consider two strategies: A means you go for it. B means you chicken out after some time. The reward for winning is b , the cost of a collision is $-c$. If both players decide to chicken out, then the chance that you will win is $1/2$. Hence the payoff matrix is

$$\begin{array}{c} \begin{array}{cc} & A & B \\ A & \begin{pmatrix} -c & b \end{pmatrix} \\ B & \begin{pmatrix} 0 & \frac{b}{2} \end{pmatrix} \end{array} \end{array} \quad (4.31)$$

Comparing the entries in each column of the matrix leads to the same conclusion as in the Hawk-Dove game. It is always best to adopt the opposite strategy as your opponent. A mixed strategy between A and B is ESS.

In the Snowdrift game, two drivers are stuck on their way home, because the road is blocked by a snowdrift. You have a choice between cooperation and defection. Cooperation means that you get out of the car and shovel. Defection means that you remain in your car, relax, listen to music, and let the other person shovel. If both cooperate, the amount of work is only half as much. If neither cooperates, each is stuck in the snowdrift until the snowplow comes along. The benefit for getting home is b , the cost of shoveling in the cold is $-c$. The payoff matrix is

$$\begin{array}{c}
 \\
 \text{C} \\
 \text{D}
 \end{array}
 \begin{array}{cc}
 \text{C} & \text{D} \\
 \left(\begin{array}{cc}
 b - \frac{c}{2} & b - c \\
 b & 0
 \end{array} \right)
 \end{array}
 \tag{4.32}$$

If $b > c$, we have the same structure as in the Hawk-Dove game. It is best to defect when the other person cooperates. It is best to cooperate when the other person defects. If $b < c$, it is best to defect no matter what the other person does. This leads us to the Prisoner's Dilemma.

4.9 GAME THEORY AND ECOLOGY

Ecologists investigate the interaction of species and how their abundance changes over time. The logistic map, which we encountered in Chapter 2, is an important equation of ecology, demonstrating that complicated time series of species abundance can be generated by very simple rules. The fundamental mathematical approach to ecology is given by the Lotka-Volterra equation. The logistic map is a one-dimensional Lotka-Volterra equation in discrete time.

4.9.1 Predator and Prey

The hostilities between the Austrian and Italian navies during World War I disrupted fishing in the Adriatic Sea. After the war, it was observed that the number of predatory fish had increased. "Why does war favor sharks?" was the question that was posed to Italy's leading physicist of the time, Vito Volterra.

Volterra wrote down the following equations. Let x and y denote, respectively, the abundance of prey and predator fish. Prey reproduce at rate ax . They are eaten by predators at rate bxy . Predators die at rate cy and reproduce at rate dx . Thus we have the following system of two nonlinear differential equations

$$\begin{aligned}
 \dot{x} &= x(a - by) \\
 \dot{y} &= y(-c + dx)
 \end{aligned}
 \tag{4.33}$$

If there are no predators, $y = 0$, then the population of prey fish grows exponentially as

$$x(t) = x(0)e^{at}. \quad (4.34)$$

If there are no prey, $x = 0$, then the population of predators declines exponentially as

$$y(t) = y(0)e^{-ct}. \quad (4.35)$$

The point $x = 0$ and $y = 0$ is a saddle point of the system. There is one interior equilibrium given by

$$x^* = c/d \quad \text{and} \quad y^* = a/b. \quad (4.36)$$

A linear stability analysis reveals that this equilibrium is neutrally stable. It is surrounded by infinitely many periodic orbits. The abundance of prey and the abundance of predators oscillate indefinitely. The oscillatory period is $2\pi/\sqrt{ac}$, but the amplitude depends on the initial condition. The time averages of the oscillations are given by the equilibrium values x^* and y^* . Neutral oscillations represent an ungeneric phenomenon, because small modifications of the differential equation (4.33) can destroy neutral stability.

Volterra could now answer the question why the reduced activity of fishing during the war had favored predators. Fishing decreases the reproductive rate of prey from a to $a - k$ and increases the death rate of predators from c to $c + m$. Hence, with fishing the time averages of the oscillations are given by

$$x_F^* = (c + m)/d \quad \text{and} \quad y_F^* = (a - k)/b. \quad (4.37)$$

The relative abundance of predators is greater in the absence of fishing:

$$y^*/x^* > y_F^*/x_F^*. \quad (4.38)$$

4.9.2 Kolmogorov's Predator-Prey Theorem

The neutral stability of equation (4.33) is a consequence of the particular assumptions for the interaction between prey and predator. There have been many extensions toward increased realism, such as including maximum car-

rying capacities for prey and predators and saturating functional responses of predator to prey. All such extensions tend to remove the unrealistic neutral stability. Instead, the abundance of predators and prey will usually converge to either a stable equilibrium or a stable limit cycle. In contrast to neutral oscillations, a stable limit cycle is a robust phenomenon. After small perturbations, the trajectory returns to the limit cycle. The period and amplitude do not depend on the initial condition, but are entirely determined by the parameters of the equation.

In 1936 the Russian mathematician A. N. Kolmogorov wrote a general predator-prey equation in the following form

$$\begin{aligned}\dot{x} &= xF(x, y) \\ \dot{y} &= yG(x, y)\end{aligned}\tag{4.39}$$

Here F and G are continuous functions with continuous first derivatives. Kolmogorov's theorem states that equation (4.39) has either a stable limit cycle or a stable equilibrium if the following conditions hold:

- (i) $\partial F/\partial y < 0$
- (ii) $x(\partial F/\partial x) + y(\partial F/\partial y) < 0$
- (iii) $\partial G/\partial y < 0$
- (iv) $x(\partial G/\partial x) + y(\partial G/\partial y) > 0$
- (v) $F(0, 0) > 0$

Furthermore, there must exist constants, $A > 0$ and $B > C > 0$, such that

- (vi) $F(0, A) = 0$
- (vii) $F(B, 0) = 0$
- (viii) $G(C, 0) = 0$

The biological interpretation of these conditions is illuminating: (i) the per capita rate of increase of prey is a decreasing function of predator abundance; (ii) for any given ratio of the two species, the rate of increase of prey is a decreasing function of population size; (iii) the per capita rate of increase of predators is a decreasing function of their abundance; (iv) for any given ratio

of the two species, the rate of increase of predators is an increasing function of population size; (v) when both populations are small, the abundance of prey can increase; (vi) there exists a predator population size sufficiently large to prevent further increase of the prey population; (vii) there exists a critical prey population size, B , beyond which it can no longer grow even if there are no predators, which means the ecosystem has a maximum carrying capacity for the prey species; (viii) there exists a critical prey population size, C , below which predators cannot grow even if they are rare.

For a thorough discussion of such issues and many additional aspects of theoretical ecology, you must seek out Robert May's classic *Stability and Complexity of Model Ecosystems*, first published in 1973.

4.9.3 The Lotka-Volterra Equation

Equation (4.33) was also studied by the American biologist Alfred Lotka in the context of chemical kinetics and, therefore, became known as Lotka-Volterra equation. The general Lotka-Volterra equation describes the interaction of n species and is of the form

$$\dot{y}_i = y_i \left(r_i + \sum_{j=1}^n b_{ij} y_j \right) \quad i = 1, \dots, n \quad (4.40)$$

The abundance of species i is given by y_i , which is a non-negative number. Hence the equation is defined on the positive orthant R_+^n , which is given by the set of all points (y_1, y_2, \dots, y_n) with $y_i \geq 0$ for all i . The growth rate of species i is given by r_i . The interaction between species i and j is given by b_{ij} . The parameters r_i and b_{ij} can be positive, zero, or negative.

It turns out that the Lotka-Volterra equation (4.40) and the replicator equation (4.21) are equivalent. A replicator equation with n strategies can be transformed into a Lotka-Volterra equation with $n - 1$ species. The $n \times n$ matrix $A = [a_{ij}]$ defines the interaction of the n strategies in the replicator equation

$$\dot{x}_i = x_i \left(\sum_{j=1}^n a_{ij} x_j - \phi \right) \quad i = 1, \dots, n \quad (4.41)$$

The Lotka-Volterra equation

$$\dot{y}_i = y_i \left(r_i + \sum_{j=1}^{n-1} b_{ij} y_j \right) \quad i = 1, \dots, n-1 \quad (4.42)$$

with the parameters $r_i = a_{in} - a_{nn}$ and $b_{ij} = a_{ij} - a_{nj}$ is equivalent to (4.41).

Let $y = \sum_{i=1}^{n-1} y_i$. The equivalence can be shown with the transformation $x_i = y_i/(1+y)$ for $i = 1, \dots, n-1$ and $x_n = 1/(1+y)$.

Therefore, whatever result holds for one equation will hold for the other. The equivalence of the Lotka-Volterra and replicator equations represents a beautiful bridge between aspects of theoretical ecology and evolutionary game theory.

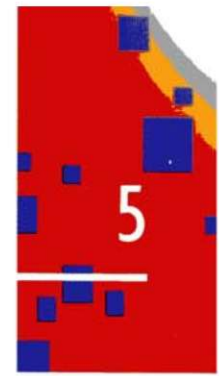
We will see that theoretical ecology is the foundation for many studies in mathematical biology. For example, viruses and cells can behave as predator and prey. Immune cells, in turn, can “prey” on infected cells. Virus dynamics in infected hosts represent a “microecology.”

SUMMARY

- ◆ Evolutionary game theory is the study of frequency-dependent selection.
- ◆ Games can be formulated in terms of a payoff matrix, which specifies the payoff for one strategy when interacting with another. Evolutionary game theory interprets payoff as fitness: successful strategies reproduce faster.
- ◆ A Nash equilibrium is a strategy with the following property: if two players adopt a Nash equilibrium, then neither player can improve her payoff by switching to another strategy.
- ◆ If an entire population (of infinite size) adopts an evolutionarily stable strategy (ESS), then no other strategy can invade.
- ◆ The replicator equation describes deterministic evolutionary game dynamics. For $n = 2$ strategies, there can be dominance, coexistence, bi-stability, or neutrality. For $n \geq 3$ strategies, there can be heteroclinic cycles. For $n \geq 4$ strategies, there can be limit cycles and chaos.

- ◆ In the Rock-Paper-Scissors game there can be either damped oscillations to a stable equilibrium or oscillations with increasing amplitude that eventually lead to the random elimination of two strategies.
- ◆ The Hawk-Dove game explains why many animals can resolve conflict without full-blown escalation.
- ◆ Chicken and Snowdrift are games that are similar to Hawk-Dove.
- ◆ The replicator equation is equivalent to the Lotka-Volterra equation of ecology. Evolutionary game theory and ecology have the same mathematical foundations.
- ◆ Kolmogorov's theorem specifies when a two-dimensional predator-prey system has a stable equilibrium or a stable limit cycle.

PRISONERS OF THE DILEMMA



TWO PEOPLE are suspected of having committed a joint crime. The suspects are confined to different rooms and cannot talk to each other. The police do not have sufficient evidence to convince a jury. The state attorney offers each of the suspects a deal: confess your crime, become a witness for the prosecution, and thereby avoid a prison sentence. If one prisoner confesses while the other does not, then the first will go free immediately; the second will receive a prison sentence of ten years. If both confess, they will each receive seven years. If neither of them confesses, then both will go free after one year because nothing can be proved. Here is the payoff matrix

$$\begin{array}{l} \text{Remain silent} \\ \text{Confess} \end{array} \begin{pmatrix} \begin{array}{cc} \text{Remain silent} & \text{Confess} \\ -1 & -10 \\ 0 & -7 \end{array} \end{pmatrix} \quad (5.1)$$

What should they do? What does this have to do with evolutionary biology, anyway?

In fact, for biology the problem is as old as evolution itself. Evolutionary progress, the construction of new features, often requires the cooperation of simpler parts that are already available. For example, replicating molecules had to cooperate to form the first cells. Single cells had to cooperate to form the first multicellular organisms. The soma cells of the body cooperate and help the cells of the germ line to reproduce. Animals cooperate to form social structures, groups, and societies. Worker bees risk their lives to defend the beehive. Moreover they forgo their own reproductive potential and instead raise the offspring of another individual, the queen. In some bird species, helpers assist the parents in feeding their young. Humans cooperate on a large scale, giving rise to cities, states, and countries. Cooperation allows specialization. Nobody needs to know everything. But cooperation is always vulnerable to exploitation by defectors.

Imagine two individuals who can either cooperate, C , or defect, D . If both cooperate, then both get 3 points. If one cooperates, while the other defects, the cooperator gets 0 points while the defector gets 5. If both defect, they get 1 point each. The payoff matrix is

$$\begin{matrix} & C & D \\ C & \begin{pmatrix} 3 & 0 \end{pmatrix} \\ D & \begin{pmatrix} 5 & 1 \end{pmatrix} \end{matrix} \tag{5.2}$$

This payoff matrix has the same structure as the situation that was presented to the prisoners. What would you do? Cooperate or defect?

You should analyze the game as follows. Assuming the other person will cooperate, you have a choice between receiving 3 points for your cooperation or 5 points for your defection. Thus if the other person cooperates, it is best for you to defect. Assuming the other person will defect, you have a choice between receiving 0 points for your cooperation or 1 point for your defection. Thus if the other person defects, it is also best for you to defect. Therefore, no matter what the other person does, it is best for you to defect.

If the other person analyzes the game in the same rational way, then both of you will come to the conclusion to defect. Both of you will end up with only 1 point. This is less than the 3 points you could have received for mutual cooperation.

In the **Prisoner's Dilemma** you can either cooperate, *C*, or defect, *D*

The payoff matrix is:

		My move	
		<i>C</i>	<i>D</i>
Your move	<i>C</i>	3 0	0 5
	<i>D</i>	5 1	1 3

↑
Your payoff

Figure 5.1 The Prisoner's Dilemma captures the essential problem of cooperation. Assuming that I will cooperate, you can get 3 points if you cooperate but 5 points if you defect. Assuming that I will defect, you can get 0 points if you cooperate but 1 point if you defect. Hence, no matter what I will do, it is better for you to defect. Defection is the "rational" (= payoff maximizing) strategy. If, however, I analyze the game in the same way that you do, then we both choose defection and both get 1 point. We could have received 3 points each, had we both chosen cooperation. But cooperation is "irrational." This is the dilemma.

This is the dilemma: rational players who act in order to maximize their payoff defect in the Prisoner's Dilemma (PD). Mutual cooperation leads to a higher payoff than mutual defection, but cooperation is irrational (Figure 5.1).

Do not take offense at the terms "rational" and "irrational" in this context. Note that the payoff describes exactly what the players want. There is no hidden agenda. The rewards (material or immaterial) are completely specified by the payoff matrix. Under this assumption, a rational player is defined as one who acts in a way so as to maximize her payoff. This is a simple and straightforward concept.

Experimental game theory, however, more often than not shows that humans do not behave rationally. They are guided by instincts that might have evolved via different situations. In the Prisoner's Dilemma, humans often try to cooperate. Only when they learn that it does not work will they switch to defection.

Returning to our prisoners, cooperation means not to cooperate with the state attorney but to cooperate with your partner and remain silent. If both of you remain silent, nothing can be proved. Defection means confession. If both of you defect, both will get a long prison sentence. No matter what your partner does, it is better for you to defect. The rational analysis suggests that both prisoners will confess and spend seven years in jail.

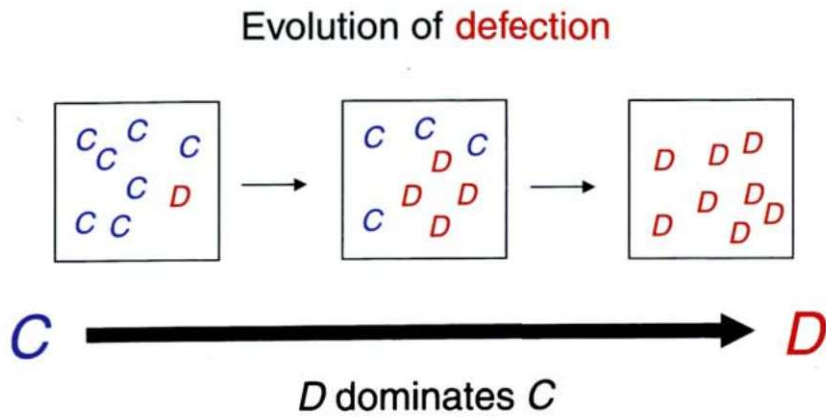


Figure 5.2 Natural selection chooses defection. In a mixed population of cooperators, C , and defectors, D , the latter always have a higher payoff. In the framework of evolutionary game theory, therefore, defectors reproduce faster and outcompete cooperators. Under natural selection, the average fitness of the population continuously declines. When the population consists only of defectors, the average fitness has reached its minimum. Specific mechanisms are needed for natural selection to favor cooperation.

If rationality leads to defection, which is a suboptimum outcome for both, then what does natural selection suggest? Imagine a population of cooperators and defectors. The frequency of cooperators is given by x , and the frequency of defectors by $1 - x$. The average payoff for cooperators is $f_C = 3x$. The average payoff for defectors is $f_D = 5x + 1 - x = 4x + 1$. Thus defectors always have a higher fitness than cooperators. Defectors dominate cooperators. Natural selection steadily increases the frequency of defectors until cooperators have become extinct. Natural selection too favors defection (Figure 5.2).

5.1 DIRECT RECIPROCITY

Consider the payoff matrix

$$\begin{array}{c}
 C \quad D \\
 \begin{array}{c} C \\ D \end{array} \begin{pmatrix} R & S \\ T & P \end{pmatrix}
 \end{array} \tag{5.3}$$

The game is a Prisoner's Dilemma if $T > R > P > S$. The temptation to defect, T , exceeds the reward for mutual cooperation, R , which is greater than the punishment, P , for mutual defection, which trumps the sucker's payoff, S . In addition, it is usually required that $R > (T + P)/2$. If this is not the case, then an agreement to alternate cooperation and defection leads to a higher payoff than pure cooperation in a repeated game.

Direct reciprocity refers to the concept that the game is not played just once but is repeated several times between the same two players. In this case, cooperation becomes a promising option. Imagine the game is repeated m times. Consider two strategies. "GRIM" cooperates on the first move and then cooperates as long as the opponent does not defect. If the opponent defects once it will switch permanently to defection. It will never forgive. The other strategy is "Always defect" (ALLD).

The payoff matrix is given as follows

$$\begin{array}{cc} & \begin{array}{cc} \text{GRIM} & \text{ALLD} \end{array} \\ \begin{array}{c} \text{GRIM} \\ \text{ALLD} \end{array} & \left(\begin{array}{cc} mR & S + (m - 1)P \\ T + (m - 1)P & mP \end{array} \right) \end{array} \quad (5.4)$$

If $mR > T + (m - 1)P$, then GRIM is a strict Nash equilibrium when competing with ALLD. In terms of a game between two players, if both follow this rule, then neither player can improve by switching to ALLD. In terms of evolutionary dynamics, if the whole population uses GRIM, then ALLD cannot invade: selection opposes ALLD at low frequency. GRIM is stable against invasion by ALLD if the number of rounds, m , exceeds a critical value,

$$m > \frac{T - P}{R - P}. \quad (5.5)$$

For the payoff values used in the matrix (5.2), we have $m > 2$.

Therefore, we seem to have found a mechanism to stabilize cooperation once it has been established. Note, however, that ALLD is also a strict Nash equilibrium, because mP is always greater than $S + (m - 1)P$. Thus we have no evolutionary mechanism to explain the emergence of cooperation. We will address this problem later.

There is an even more acute problem to deal with. Imagine we both know that we are playing the game for m rounds. There is no incentive for you to cooperate in the last round. For the last round, the same analysis applies as for the nonrepeated Prisoner's Dilemma. Realizing this point, we might choose to play GRIM with one small modification: we will certainly defect in the last round. Let us denote this strategy by GRIM*. The payoff matrix for GRIM versus GRIM* is given by

$$\begin{array}{c} \text{GRIM} \\ \text{GRIM}^* \end{array} \begin{pmatrix} \begin{array}{cc} \text{GRIM} & \text{GRIM}^* \\ mR & (m-1)R + S \\ (m-1)R + T & (m-1)R + P \end{array} \end{pmatrix} \quad (5.6)$$

We notice that GRIM is dominated by GRIM*. A population of GRIM players can be invaded by a small fraction of GRIM*. Once everybody plays GRIM*, the same argument as before applies to the second-to-last round. It is not rational to cooperate in the penultimate round if the ultimate round is mutual defection. The argument can be repeated back to the first round. We can thus write down a sequence of strategies starting with GRIM that is dominated by the variant strategy that defects in the last round, which is dominated by the strategy that defects in the last two rounds and so on until we come to ALLD, which is a strict Nash equilibrium. In this strategy space, the only strict Nash equilibrium and the only evolutionarily stable strategy is ALLD.

We note again that humans do not use this backward induction in experimental situations. People might realize that it is best to defect in the end, but they do not carry the argument rigorously to its logical conclusion. Why not? One explanation is that the human strategic instincts are not formed by playing games with a well-defined number of rounds. Imagine that it is not quite known when the interaction will end. There might always be another round. There is always a tomorrow in our plans. History has a beginning, but it has no end.

5.1.1 Tomorrow Never Dies

Let us therefore study a repeated Prisoner's Dilemma with a variable, rather than a fixed number of rounds. Suppose that after each round there is a prob-

ability w that another round will be played. Hence the expected number of rounds is $\bar{m} = 1/(1 - w)$. The payoff matrix between GRIM and ALLD becomes

$$\begin{array}{cc} & \begin{array}{cc} \text{GRIM} & \text{ALLD} \end{array} \\ \begin{array}{c} \text{GRIM} \\ \text{ALLD} \end{array} & \left(\begin{array}{cc} \bar{m}R & S + (\bar{m} - 1)P \\ T + (\bar{m} - 1)P & \bar{m}P \end{array} \right) \end{array} \quad (5.7)$$

GRIM is evolutionary stable if $\bar{m} > (T - P)/(R - P)$. Nothing changes except now there is no strategy that can simply defect on the last move.

But is GRIM an ideal strategy for the repeated Prisoner's Dilemma? What if the opponent defects only once, perhaps to see if he can get away with it? Is it best never to cooperate with such an individual again? A strategy with some mechanism of reconciliation might achieve a higher payoff than GRIM.

What is the set of all strategies for the repeated Prisoner's Dilemma? A deterministic strategy is a rule that assigns to any history of the game the decision to cooperate or to defect in the next move. A stochastic strategy is a rule that assigns to any history of the game the probability of cooperating (or defecting) in the next move. Each round of the game has four possible outcomes: CC , CD , DC , DD . There are $2^4 = 16$ deterministic strategies that consider only the last round of the game; for each of the four possible outcomes the strategy prescribes the decision to cooperate or to defect. Hence a deterministic strategy with a memory of one move can be encoded by a binary string of length 4. For example, 0000 means "always defect," while 1000 means cooperate only after a CC and defect otherwise. There are 2^{16} deterministic strategies that consider the last two moves of the game. There are 2^{4^m} -deterministic strategies that consider the last m moves. Stochastic strategies that consider the last m moves of the game span a 4^m dimensional strategy space. Each dimension is given by the interval $[0, 1]$, denoting a probability. For a game of arbitrary length, the space of all possible strategies has infinitely many dimensions. It is impossible to set up a computer simulation that considers every possible strategy in the repeated PD. How can we find successful strategies for direct reciprocity?

5.2 AXELROD'S TOURNAMENTS

The search for an optimum strategy led Robert Axelrod, a political scientist at the University of Michigan at Ann Arbor, to a brilliant idea. In 1978 he decided to conduct a Prisoner's Dilemma championship. He invited people from all over the world to submit strategies for the repeated game. The strategies had to be formulated in terms of computer programs. Every strategy played every other strategy. The payoffs from each encounter were added up. In the end, Axelrod analyzed which strategy had the highest total payoff.

There were fourteen contestants. Some programs were clever devices to deceive the opponent or to predict the opponent's behavior. The winning strategy was the simplest of all. It was Tit-for-tat (TFT). TFT starts with a cooperation and then does whatever the opponent did in the previous round. TFT will answer C for C and D for D. Playing against TFT is like playing the mirror image of yourself shifted by one round. TFT was submitted by the well-known game theorist Anatol Rapoport.

Axelrod published the results together with a comprehensive analysis. Then he invited people to submit strategies for a second world championship. This time sixty-three strategies were sent in. Many people submitted strategies that would have won the first tournament. Among these was Tit-for-two-tats submitted by John Maynard Smith. Only Rapoport fielded TFT. And it won again.

While it is possible to calculate the optimum strategy against a known set of contestants, it is difficult to predict the best strategy for an unknown set of contestants. Axelrod published a book, *The Evolution of Cooperation*, which became a best-seller. TFT was the unquestioned world champion in the repeated PD.

Axelrod praises TFT in his book for having a collection of important properties. It is "nice," which means it is never the first to defect. TFT never tries to get more than the opponent in a direct pairwise comparison. In each single tournament match, it will at most have received the same number of points as its opponent, but never more. The payoff sum over all matches, however, was higher for TFT than for any other strategy. Hence, TFT is not triumphant by directly beating other strategies. Its success lies in the fact that on average it gets a higher payoff from strategy X than other strategies get from strategy X . TFT is apparently very successful at inducing cooperation from other strategies.

Moreover, TFT is stable against invasion by ALLD if the average number of rounds, \bar{m} , is not too small. TFT playing ALLD will cooperate in the first round, but defect in every subsequent round. The payoff matrix for TFT versus ALLD is the same as for GRIM versus ALLD

$$\begin{array}{cc} & \begin{array}{cc} \text{TFT} & \text{ALLD} \end{array} \\ \begin{array}{c} \text{TFT} \\ \text{ALLD} \end{array} & \left(\begin{array}{cc} \bar{m}R & S + (\bar{m} - 1)P \\ T + (\bar{m} - 1)P & \bar{m}P \end{array} \right) \end{array} \quad (5.8)$$

TFT can resist invasion by ALLD if $\bar{m} > (T - P)/(R - P)$.

TFT has the advantage over GRIM that it can resume cooperation if the opponent cooperates. Unlike GRIM, TFT is not locked into permanent defection.

5.2.1 Achilles' Heel

Axelrod's original tournaments were conducted in an error-free digital universe, but real-world situations are permeated by mistakes. A "trembling hand" can lead to a misimplementation of one's own action. A "fuzzy mind" can cause the misinterpretation of the opponent's move.

It turns out that TFT has an Achilles' heel (Figure 5.3). In the presence of mistakes, two TFT players achieve a very low payoff. A single mistake moves the game from mutual cooperation to alternating between cooperation and defection. A second mistake can lead to mutual defection. In the long run, two TFT players, with a small chance of making mistakes, obtain the same payoff as two random players tossing coins to decide between C and D in every move. The payoff for two TFT players in the presence of small amounts of behavioral noise is given by

$$A(TFT, TFT) = \frac{R + T + P + S}{4}. \quad (5.9)$$

This payoff is certainly less than R , because $R > (T + S)/2$ and $R > P$. Hence TFT is weak in the presence of mistakes. As we shall soon see, mistakes imply that TFT can be invaded and even dominated by many other strategies.

Tit-for-tat cannot correct mistakes

TFT: C C C ^{*}D C D C D D D ...
TFT: C C C C D C D ^{*}D D D ...

Figure 5.3 The Achilles' heel of a world champion: Tit-for-tat cannot correct mistakes. If an error (red asterisk) occurs, then the game switches from mutual cooperation to alternating cooperation and defection. Another error can bring the game to mutual defection. Further errors lead back to cooperation. But in the long run, the expected payoff between two TFT players is the same as for two random players who flip coins to determine whether to cooperate or to defect. Errors reduce the performance of TFT.

There is another weakness even in the absence of mistakes. Imagine TFT playing "Always cooperate" (ALLC). The payoff matrix is given by

$$\begin{array}{cc} & \begin{array}{cc} \text{TFT} & \text{ALLC} \end{array} \\ \begin{array}{c} \text{TFT} \\ \text{ALLC} \end{array} & \begin{pmatrix} \bar{m}R & \bar{m}R \\ \bar{m}R & \bar{m}R \end{pmatrix} \end{array} \quad (5.10)$$

Both players cooperate on every move. Therefore TFT is neither a strict Nash equilibrium nor an evolutionarily stable strategy. In a finite population, random drift can lead from TFT to ALLC (Figures 5.4 and 5.5).

Axelrod's world series had no noise in terms of strategical mistakes or random drift. Could these effects lead to the downfall of the world champion, TFT?

5.3 REACTIVE STRATEGIES

As we have seen, the strategy space of the repeated PD is enormous. Any particular calculation can only explore a limited region of this space. Let us now consider the set of "reactive strategies." These strategies are given by two parameters: p denotes the probability of cooperating if the opponent has cooperated in the previous move; q denotes the probability of cooperating if

The repeated Prisoner's Dilemma

ALLC: *C C C C C C C ...*

ALLD: *D D D D D D D ...*

ALLD: *D D D D D D D ...*

TFT: *C D D D D D D ...*

TFT: *C C C C C C C ...*

TFT: *C C C C C C C ...*

Figure 5.4 Three simple strategies in the repeated Prisoner's Dilemma: Always cooperate (ALLC), Always defect (ALLD), and Tit-for-tat (TFT). ALLC is exploited by ALLD. ALLD can exploit TFT only on the first move; then TFT switches to defection. Two TFT players, in contrast, can cooperate on every move (if no errors occur). Therefore, in a game between ALLD and TFT, ALLD receives a slightly higher payoff than TFT, but two TFT players receive a much higher payoff still. TFT is successful for this reason.

ALLD dominates ALLC:

ALLD ← ALLC

ALLD and TFT are bistable:

ALLD ← ○ → TFT

ALLC and TFT are neutral:

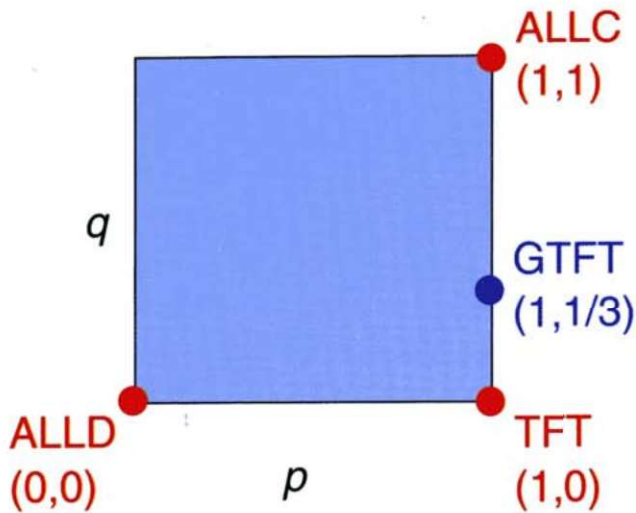
ALLC — TFT

Figure 5.5 Pairwise selection dynamics of the three basic strategies. In a mixed population of ALLD and ALLC, the former always has a higher fitness. In a mixed population of ALLD and TFT, ALLD has a higher fitness only when TFT is rare; otherwise TFT has a higher fitness. Selection dynamics are bistable. There is an unstable equilibrium (red circle) that acts as an invasion barrier: if the frequency of TFT is greater than this unstable equilibrium, then TFT can eliminate ALLD. In a mixed population of ALLC and TFT, all players have the same fitness. Therefore TFT is not evolutionarily stable.

the opponent has defected in the previous move. Note that reactive strategies have a very short memory: they only take into account what the opponent has done during the previous round. They do not even consider their own previous move.

A particular reactive strategy, $S(p, q)$, is a point in the unit square. Three corners of this square represent familiar strategies: (i) $S(0, 0)$ denotes ALLD; (ii) $S(1, 1)$ denotes ALLC; and (iii) $S(1, 0)$ denotes TFT. The fourth corner,

Reactive strategies



p : Probability of cooperating after opponent's C

q : Probability of cooperating after opponent's D

Figure 5.6 Reactive strategies are defined by two parameters: p denotes the probability of cooperating, if the opponent has cooperated in the previous round, while q denotes the probability of cooperating, if the opponent has defected in the previous round. A particular reactive strategy is given by a pair (p, q) . Examples of reactive strategies include: ALLD $(0, 0)$, ALLC $(1, 1)$, and TFT $(1, 0)$. A TFT player with a 10% error rate is given by $(0.9, 0.1)$. It can be shown that TFT is the best catalyst for the emergence of cooperation, but once TFT has established cooperation it is rapidly replaced by Generous Tit-for-tat (GTFT). For the standard payoff values, $R = 3$, $T = 5$, $P = 1$, and $S = 0$, GTFT is given by $(1, 1/3)$. GTFT forgives on average every third defection. Natural selection among reactive strategies leads to the emergence of forgiveness.

$S(0, 1)$, represents the paradoxical strategy that cooperates when the opponent has defected and defects when the opponent has cooperated (Figure 5.6).

Can we derive a complete understanding of evolutionary dynamics within the set of reactive strategies? The repeated PD between two such strategies is a Markov chain on the state space (CC, CD, DC, DD) . Let us label these states 1 to 4. State 1 means we both cooperate, CC . State 2 means I cooperate but you defect, CD . State 3 means I defect while you cooperate, DC . State 4 means we both defect, DD .

My reactive strategy is given by $S_1(p_1, q_1)$, while yours is $S_2(p_2, q_2)$. The Markov chain is defined by a 4×4 transition matrix, $M = [m_{ij}]$. The entry m_{ij} denotes the probability of moving from state i to state j . We have

$$M = \begin{array}{c} \\ CC \\ CD \\ DC \\ DD \end{array} \begin{array}{cccc} CC & CD & DC & DD \\ \left(\begin{array}{cccc} p_1 p_2 & p_1(1-p_2) & (1-p_1)p_2 & (1-p_1)(1-p_2) \\ q_1 p_2 & q_1(1-p_2) & (1-q_1)p_2 & (1-q_1)(1-p_2) \\ p_1 q_2 & p_1(1-q_2) & (1-p_1)q_2 & (1-p_1)(1-q_2) \\ q_1 q_2 & q_1(1-q_2) & (1-q_1)q_2 & (1-q_1)(1-q_2) \end{array} \right) \end{array} \quad (5.11)$$

For example, the transition probability from state CD to state DC is given by $(1 - q_1)p_2$. This term is the product of two probabilities: $(1 - q_1)$ denotes the probability that I will defect given that you have defected in the previous round, while p_2 is the probability that you will cooperate given that I have cooperated in the previous round. All entries of the M matrix can be constructed in this way.

Let \vec{x}_t denote the probability distribution of the game after t rounds. This means the vector \vec{x}_t has four components that determine the probabilities that the game is in the four possible states, (CC, CD, DC, DD) . We obtain the probability distribution after the next time step by multiplying the vector with the transition matrix

$$\vec{x}_{t+1} = \vec{x}_t M. \quad (5.12)$$

A stochastic matrix is regular if there exists a positive number k such that all elements of the matrix M^k are positive. If the matrix M is regular, then there exists a unique left-hand eigenvector, \vec{x} , associated with the eigenvalue one,

$$\vec{x} = \vec{x} M. \quad (5.13)$$

The components of this normalized eigenvector represent the stationary distribution of the Markov chain.

Let us introduce the quantities $r_1 = p_1 - q_1$ and $r_2 = p_2 - q_2$. The matrix M is regular if and only if $|r_1 r_2| < 1$. Because of the special symmetry of the M

matrix, the stationary distribution is of the form

$$\vec{x} = [s_1s_2, s_1(1 - s_2), (1 - s_1)s_2, (1 - s_1)(1 - s_2)] \quad (5.14)$$

where

$$s_1 = \frac{q_2r_1 + q_1}{1 - r_1r_2} \quad (5.15)$$

and

$$s_2 = \frac{q_1r_2 + q_2}{1 - r_1r_2}. \quad (5.16)$$

Note that s_1 and s_2 represent the probabilities that players 1 and 2 cooperate in the stationary distribution. The expected payoff for strategy S_1 versus S_2 is given by

$$\begin{aligned} E(S_1, S_2) \\ = Rs_1s_2 + Ss_1(1 - s_2) + T(1 - s_1)s_2 + P(1 - s_1)(1 - s_2). \end{aligned} \quad (5.17)$$

Equations (5.15–5.17) provide the payoff for the game between any two strategies as long as $|r_1r_2| < 1$. All the cases with $|r_1r_2| = 1$ include deterministic strategies, for which the payoff can be easily calculated.

5.3.1 “In Silico” Evolution

Let us perform an experiment to understand the evolutionary dynamics of reactive strategies. Use a random number generator that returns a uniform distribution on $[0, 1]$ to produce $n = 100$ reactive strategies. Let us use the same payoff values that Axelrod chose for his two tournaments: $R = 3$, $T = 5$, $S = 0$, and $P = 1$. Calculate the $n \times n$ payoff matrix using equation (5.17). Insert the matrix into the replicator equation. Assume that all n strategies are equally abundant at time $t = 0$. Observe the evolutionary trajectory.

In most cases, you will make the following observation. Many strategies converge to extinction. The most cooperative strategies disappear first. After some time, only a single strategy is left. This strategy is the one that is closest to ALLD. You have witnessed yet again the evolution of defection. You have also seen the destructive potential of frequency-dependent selection: the average

Generous tit-for-tat can correct mistakes

GTFT: C C C ^{*}D C D C C C C . . .
GTFT: C C C C D C C C C . . .

Figure 5.7 In contrast to TFT, GTFT can correct mistakes. With a certain probability, a sequence of cooperation and defection leads back to mutual cooperation. The expected payoff for two GTFT players is higher than for two TFT players.

fitness of the population declined steadily from about $(R + T + S + P)/4 = 9/4$ to $P = 1$.

In some cases, however, something remarkable happens. Suppose the initial ensemble contains, by chance or by design, a strategy that is close to TFT. Initially there is again an evolution toward ALLD. But when almost all strategies have disappeared, the embattled minority of TFT-like individuals fight the mighty ALLD. The reciprocators' frequency rises suddenly, while that of the defectors declines. But the reign of TFT is short-lived. Its moment of glory is brief and fleeting. TFT is soon replaced by strategies that are close to $p = 1$ and $q = 1/3$. Here the selective dynamics end.

Who is the new champion that has appeared on the cooperative wave initiated by TFT? The strategy $p = 1$ and $q = 1/3$ cooperates whenever the opponent has cooperated, but also cooperates one out of three times when the opponent has defected. We call this strategy "Generous Tit-for-tat" (GTFT). It is more forgiving than strict TFT when faced with a defection. This has two consequences.

(i) When one GTFT individual plays another, each receives an average payoff per round that is very close to the full reward for mutual cooperation, R . In contrast, two TFT players only obtain $(R + P + T + S)/4$. GTFT can make up for mistakes (Figure 5.7).

(ii) When GTFT plays ALLD, it has a slightly lower payoff than when TFT plays ALLD. Thus GTFT is less able to swing the tide away from defection than TFT is. Hence we need TFT to initiate cooperation. But once a cooperative population has been established, TFT will be replaced by GTFT.

5.4 GENEROUS TIT-FOR-TAT

For the general parameter values of the Prisoner's Dilemma, GTFT is defined as follows

$$\text{GTFT: } p = 1 \quad q = \min \left\{ 1 - \frac{T - R}{R - S}, \frac{R - P}{T - P} \right\} \quad (5.18)$$

This is the highest level of forgiveness, q , that is still resistant against invasion by ALLD. Increasing q further would allow ALLD to dominate GTFT.

We can use equation (5.17) to show that GTFT is optimum in the following sense: among all reactive strategies that can resist invasion by ALLD, it affords the highest payoff for a population adopting it.

5.5 WIN-STAY, LOSE-SHIFT

Let us now consider the set of all stochastic strategies that base their decision whether to cooperate or to defect on the previous move of the game (including both the opponent's action and one's own action). Each strategy is thus defined by the conditional probabilities (p_1, p_2, p_3, p_4) to cooperate, given that the outcome of the last round was CC , CD , DC , or DD . Again the game between two such strategies, $S(p_1, p_2, p_3, p_4)$ and $S'(p'_1, p'_2, p'_3, p'_4)$, can be formulated as a Markov chain. The transition matrix of this process is given by

$$M = \begin{matrix} & \begin{matrix} CC & CD & DC & DD \end{matrix} \\ \begin{matrix} CC \\ CD \\ DC \\ DD \end{matrix} & \left(\begin{array}{cccc} p_1 p'_1 & p_1(1 - p'_1) & (1 - p_1)p'_1 & (1 - p_1)(1 - p'_1) \\ p_2 p'_2 & p_2(1 - p'_2) & (1 - p_2)p'_2 & (1 - p_2)(1 - p'_2) \\ p_3 p'_3 & p_3(1 - p'_3) & (1 - p_3)p'_3 & (1 - p_3)(1 - p'_3) \\ p_4 p'_4 & p_4(1 - p'_4) & (1 - p_4)p'_4 & (1 - p_4)(1 - p'_4) \end{array} \right) \end{matrix} \quad (5.19)$$

M is a stochastic matrix. If M is regular, then it has a unique left-hand eigenvector, $\vec{x} = (x_1, x_2, x_3, x_4)$, associated with eigenvalue one,

$$\vec{x} = \vec{x}M. \quad (5.20)$$

This eigenvector represents the stationary distribution of the Markov chain. Thus x_1 , x_2 , x_3 , and x_4 are the probabilities to be in state CC , CD , DC , and DD , respectively, after a large number of rounds. Hence the expected payoff for strategy S playing against S' is given by

$$E(S, S') = Rx_1 + Sx_2 + Tx_3 + Px_4. \quad (5.21)$$

Strictly speaking, this is the expected payoff per round when playing an infinitely repeated Prisoner's Dilemma. All of this is exactly as in section 5.3, but the transition matrix M no longer has the nice symmetry properties and hence we cannot, in general, write down a simple formula for the eigenvector \vec{x} .

The new strategy space is given by the four-dimensional cube $[0, 1]^4$. Reactive strategies are a subset of memory-one strategies fulfilling the constraints $p_1 = p_3$ and $p_2 = p_4$. ALLD is given by $(0, 0, 0, 0)$. ALLC is given by $(1, 1, 1, 1)$. TFT is given by $(1, 0, 1, 0)$. For Axelrod's payoff values, $R = 3$, $S = 0$, $T = 5$, and $P = 1$, which we will use now, GTFT is $(1, 1/3, 1, 1/3)$.

We performed a slightly different in-silico evolution experiment, adjusted to the much larger strategy space. We start with a homogeneous population using the random strategy given by $(1/2, 1/2, 1/2, 1/2)$. Every 100 generations, on average, a small amount of a new strategy is introduced. Selective dynamics are calculated using the replicator equation. The new strategy could become extinct, take over the whole population, or establish an equilibrium with the resident strategy. After some time, another random mutant is added. The new mutant is taken from a fixed random distribution over the whole strategy space. It is beneficial to use a distribution which has some bias toward the boundary of the hypercube, because the most relevant strategies live at the boundary. The dynamics differ from the classical replicator in that new mutants are added over time.

Originally this experiment was performed to confirm the success of GTFT. Indeed for some instantiations of this evolutionary process, GTFT is the ultimate winner. But more often we discovered—unexpectedly—that evolution came up with another strategy. The deterministic form of this strategy is given by $(1, 0, 0, 1)$. The strategy thus cooperates if the previous move was CC or DD . The strategy defects if the previous move was CD or DC . What does this mean?

Win-stay, lose-shift

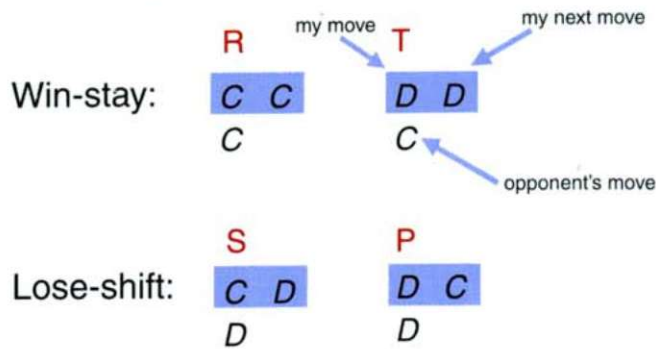


Figure 5.8 Win-stay, lose-shift (WSLS) simply monitors its own payoff: “if I am doing well, I will continue with the present action; if I am not doing well, I will try something else.” The payoff R is obtained for mutual cooperation, CC . In this case WSLS will cooperate again. The even larger payoff T is obtained for DC . In this case WSLS will defect again. The payoffs R and T are considered a “win” and therefore WSLS “stays” with its current move. The payoff S is obtained for CD . In this case WSLS will shift from C to D . The payoff P is obtained for mutual defection, DD . In this case WSLS will shift from D to C . The payoffs S and P are considered a “loss” and therefore WSLS “shifts” to another move. It is remarkable that this simple and fundamental rule can outperform both TFT and GTFT in the repeated Prisoner’s Dilemma.

Note that the strategy repeats its previous move whenever it has received a high payoff, T or R . The strategy shifts to the opposite move whenever it has received a low payoff, P or S . Thus the strategy follows the simple principle of Win-stay, lose-shift, WSLS (Figure 5.8).

Rapoport did consider WSLS, but noted that it attempts to cooperate with ALLD in every other round,

$$\begin{aligned} \text{WSLS: } & CDCDCDCD \dots \\ \text{ALLD: } & DDDDDDDD \end{aligned} \tag{5.22}$$

Therefore he called it Simpleton.

Note however that WSLS is stable against invasion by ALLD if $R > (T + P)/2$. By coincidence, Axelrod’s values have the ungeneric property that $R = (T + P)/2$. In this case, a variant of WSLS with $(1, 0, 0, 1 - \epsilon)$ is stable against invasion by ALLD. For $R < (T + P)/2$ we find that $(1, 0, 0, x)$ with $x < (R - P)/(T - R)$ is stable against invasion by ALLD.

Win-stay, lose-shift can correct mistakes

WSLS: $C C C \overset{*}{D} D C C C \dots$
WSLS: $C C C C D C C C \dots$

Win-stay, lose-shift dominates ALLC

WSLS: $C C C \overset{*}{D} D D D D \dots$
ALLC: $C C C C C C C C \dots$

Figure 5.9 Win-stay, lose-shift (WSLS) can correct mistakes. In a game between two WSLS players, a mistake (red asterisk) leads to one round of mutual defection, after which cooperation is resumed. WSLS dominates ALLC. After a mistake, WSLS switches to defection. Therefore WSLS leads to a more stable cooperation than TFT or GTFT. In contrast to WSLS, (i) TFT cannot correct mistakes and (ii) neither TFT nor GTFT is stable against random drift to ALLC.

WSLS has the nice property that it can correct occasional mistakes. If two WSLS strategists play each other, they will start with a sequence of cooperative moves and continue to cooperate until one of them defects by mistake. In the next round, both defect. The player who defected first received a very high payoff, was happy and defected again. The player who cooperated received a low payoff, was unhappy, and switched from cooperation to defection. Now both players defect, are unhappy, and switch back to cooperation. The sequence of moves is the following

$$\begin{aligned} \text{WSLS: } & C C C C C \overset{*}{D} D C C C \dots & (5.23) \\ \text{WSLS: } & C C C C C D C C C \end{aligned}$$

The asterisk indicates the erroneous move. The sequence of events seems compatible with human psychology: a mistake leads to a brief argument, after which friendship is resumed. WSLS has the advantage over TFT, because it can correct occasional mistakes. WSLS is a deterministic corrector, while GTFT is a stochastic one (Figure 5.9).

WSLS has an advantage over GTFT, too. Consider the game between WSLS and ALLC. Initially both cooperate, but after some time WSLS recognizes that

ALLC does not reciprocate defection. WSLS will then switch from mutual cooperation to exploitation. The sequence of moves is

$$\begin{aligned} \text{WSLS: } & CCCCC^*DDDDD \dots & (5.24) \\ \text{ALLC: } & CCCCCCCCCC \end{aligned}$$

Again the asterisk indicates an erroneous move. Sadly, this behavior is also compatible with human psychology: unconditional (defenseless) cooperators tend to be exploited.

In summary, TFT is the best initiator of cooperation in a society that consists largely of defectors. Once cooperation has been established, however, TFT's unforgiving retaliation causes its downfall. Forgiving strategies like GTFT or WSLS take over. A GTFT population can drift to unconditional cooperation, which in turn invites defectors. A population using WSLS is stable against neutral drift to ALLC and can resist invasion by ALLD. WSLS is no simpleton (Figure 5.10).

SUMMARY

- ◆ Cooperation means that the donor pays a cost and the recipient gets a benefit.
- ◆ In evolutionary biology, cost and benefit are measured in terms of reproductive success.
- ◆ It is not obvious how natural selection can lead from competition to cooperation.
- ◆ Yet cooperation is abundant in biology.
- ◆ The Prisoner's Dilemma is a game that captures the essence of cooperation.
- ◆ In the Prisoner's Dilemma, cooperation is dominated by defection.
- ◆ The repeated Prisoner's Dilemma is a tool for studying direct reciprocity, which presents a mechanism for the evolution of cooperation.
- ◆ Tit-for-tat is a simple but successful strategy of direct reciprocity. Tit-for-tat cooperates on the first move and then does whatever the opponent did in the previous round.

Oscillations of cooperation and defection

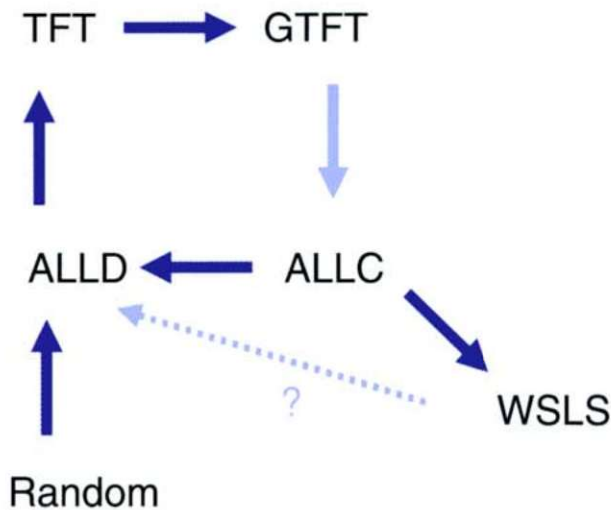


Figure 5.10 War and peace. Out of an initial “chaos” of random strategies, ALLD emerges as the first winner. Defection is born out of chaos. TFT is the best catalyst for initiating a cooperative society; only a small cluster of TFT players is needed to invade and replace ALLD. Once TFT is abundant, it is not here to stay. It is harmed by its own unforgiveness and replaced by GTFT. If everybody is nice and generous, the biological trait of retaliation is lost. ALLC emerges as a neutral mutant and takes over by random drift. A population of unconditional cooperators (ALLC) invites the rise of defectors (ALLD). There are ongoing oscillations of cooperative and defective societies in a struggle of war and peace. The cycles are interrupted by the emergence of WSLS, which can dominate ALLC and resist invasion by ALLD. Occasionally we observe evolutionary trajectories that dislodge WSLS (dotted arrow); the mechanism of this evolution is not yet fully understood.

- ◆ Tit-for-tat has two weaknesses: (i) it cannot correct mistakes; (ii) it cannot prevent neutral drift leading to “Always cooperate,” ALLC.
- ◆ Generous Tit-for-tat cooperates whenever the opponent has cooperated, and sometimes even cooperates when the opponent has defected. Generous Tit-for-tat can correct mistakes.
- ◆ Evolutionary analysis of “reactive strategies” reveals that Tit-for-tat is a catalyst for the emergence of cooperation, but not the ultimate goal. It is replaced by Generous Tit-for-tat.
- ◆ Both strategies are outcompeted by Win-stay, lose-shift, which can correct mistakes and is stable against neutral drift to ALLC.

