5 C *The Evolution of Cooperation*

Hiroo Onoda (小野田寛郎) was born in 1922 in the Empire of Japan. As a boy, he herded cattle. In 1940 he joined the army. In 1944, he traveled to a small island in the Philippines with orders to defend the island against the United States. The Empire of Japan surrendered the next year, but no one could find Onoda to tell him the bad news.

Onoda surrendered 29 years later. His 29 years of persistence had no effect on the war. No effect aside from his own long deprivation. But no one saw it that way. He returned to Japan a hero, a symbol of dedication, honor, and altruism.

Human societies make role models of people like Onoda, because they recognize the conflict between the interests of individuals and groups. The groups are not always so large as the Empire of Japan. And the individuals are not always so capable as Hiroo Onoda. But conflict is the context in which cooperation happens. Typically each individual also has ways to achieve its goals that do not promote the goals of others. This creates conflict between the interests of individuals and the interests of the groups and societies they are members of.

In this chapter, you'll meet simple game theoretic models of cooperation. You'll learn how to analyze these models to understand precisely how to express the conflict between individuals and groups. And then you'll study a general mechanism that helps cooperation to evolve. In later chapters, the problem of cooperation continues. But it gets much more complex.

Pure conflict: The prisoner's dilemma

Individual costs can undermine the production of common benefits.

Cooperation is a messy category. So it's useful to study a simple, extreme example. Then we'll consider less simple, less extreme cases.

Suppose pairs of individuals share a common territory. Each individual can choose to patrol the territory. If either individual patrols, a common benefit b is produced for both individuals. If only one individual patrols, it costs that individual c units of fitness. If both patrol, they split the cost. Each pays only $c/2$. This is the payoff matrix for this game:

	Rest	Patrol
Rest	0	b
Patrol	$b-c$	$b-c/2$

Now analyze the game's evolutionary dynamics. Patrol is evolutionarily stable if $b-c/2 > b$, which is never true when $c > 0$. A rare Rest can always invade. Rest is evolutionarily stable if $0 > b - c$, which implies $c > b$. Suppose $b > c$. Then this game has dynamics like the Hawk-Dove game: there will be a stable mix of Rest and Patrol at equilibrium. To be precise, let p be the proportion of Patrol. Then the equilibrium is at:

$$p(b - c/2) + (1 - p)(b - c) = pb + (1 - p)0$$
$$p = \frac{b - c}{b - c/2}$$

But now suppose $c > b$. If the cost of patrolling alone is greater than the benefit, no one will patrol. The population does best when someone patrols. But evolution won't sustain it.

When $c > b$, this game is an example of the **prisoner's dilemma**. In a prisoner's dilemma, non-cooperation is favored, no matter how common cooperation is in the population. Biologists and social scientists study the prisoner's dilemma a lot. Maybe too much. But they study it because it is an extreme example of the basic puzzle of cooperation: How can animals sustain cooperation when costs are private while benefits are shared? In a prisoner's dilemma, the interests of the individual and group are completely opposed. Each individual does better by resting. But the group does better when someone cooperates (patrols).

This dilemma may seem obvious to you. But there were many scholars in both biology and the social sciences who did not realize it. In the early 20th century, is was common for biologists to think that animals were designed for the good of their social groups and for social scientists to believe that people cooperate just because cooperation produces shared benefits.[1] Simple games like the prisoner's dilemma taught us that it isn't so simple.

Partial conflict: Coordination

The interests of individuals and groups are not always perfectly in opposition. But risk and mis-coordination are other reasons it can be hard to achieve cooperation.

Suppose a pair of animals who can choose to search for food that is easy but low quality or instead risky but high quality. The easy option always produces a payoff b, regardless of what the other individual does. But the risky option only

[1]For example, the framing of Mancur Olson's *The Logic of Collective Action* (1965) seems weird now. But Olson was writing for an audience that had not thought deeply about the conflict between individual and group interests.

produces $B > b$ when both individuals choose it. Otherwise it returns on average zero. This is the payoff matrix:

	Safe	Risky
Safe	b	b
Risky	0	B

The population will be better off if everyone pursues the risky option. Now what happens when Safe is common? Can Risky invade? No, because $b > 0$. Can Risky be stable when common? Yes, because $B > b$. So Risky is evolutionarily stable, but it cannot invade. This is a different kind of conflict than in the prisoner's dilemma, but just as serious. Everyone would be better off, if they could agree to choose the risky option. But there is no path for evolution to take, in this simple example. And since Safe always produces b, it is hard to evolve away from. The game above is a classic, and it is usually called the **stag hunt**.

Another kind of cooperation dilemma is **coordination**. Suppose there are two behaviors, and to receive a non-zero payoff, an individual must match the behavior of its partner. But the different behaviors produce different payoffs. Like this:

	A	B
A	b	0
B	0	B

Again assume $B > b$. The most important thing in this game is to match your partner. But both behaviors are evolutionarily stable. So if a population is fixed on A, evolution cannot evolve to B, even though it would make everyone better off. This is another type of cooperation conflict, one in which the dilemma is how to get everyone to coordinate to choose the better option.

In both of the games above, both options are evolutionarily stable. This implies that there is some proportion of each strategy at which the fitness of each is equal, an equilibrium. But this equilibrium will be unstable. If the population moves to one side or the other, natural selection will push towards an extreme of one behavior or the other dominating the population. This is the opposite dynamic of what we saw in the previous chapter, with Hawk-Dove. In Hawk-Dove, natural selection pushed the population to a stable mix. In these games, natural selection will instead push the population to the extremes.

Still the equilibrium is interesting, because it defines the **basin of attraction** for each strategy. A basin of attraction is the range of proportions that favor a strategy. For the stag hunt, we find the unstable equilibrium that defines the basins of attraction by setting the average fitness of both strategies equal and solving for the proportions that make the equality true. Let p be the proportion choosing Risky. This implies the equilibrium is found where:

$$pB + (1 - p)0 = b$$

This is true only when $p = b/B$. When $p < b/B$, natural selection favors the safe option. When $p > b/B$, natural selection favors the risky option. So if b is small enough, then it might not be so hard to for other forces to push a population over the equilibrium so it can reach the fitness-enhancing behavior. But if for example $b = 1$ and $B = 2$, then the unstable point is right in the middle and populations likely get stuck at the inferior behavior, even though the alternative is twice as good.

Can you find the unstable equilibrium in the coordination game?

Positive assortment

A general solution to the problem of evolving cooperation is when cooperative strategies tend to encounter one another more often than chance.

Think again of the patrol game. Suppose pairs were always comprised of the same strategy. So pairs of Rest and pairs of Patrol, but no mixed pairs. So Rest always gets zero, and Patrol always gets $b-c/2$. Cooperation would always evolve. If evolution can find a way for cooperative strategies to **positively assort**, this would help those strategies to evolve.

There are several mechanisms that biologists and social scientists think can produce positive assortment of cooperative strategies. We'll model some of them in later chapters. For now, let's think of assortment in purely statistical terms, without specifying the mechanism that produces it. Specifically, suppose the probability that an individual with the strategy Patrol is paired with another Patrol is:

$$\Pr(\text{Patrol}|\text{Patrol}) = r + (1-r)p$$

where r is the probability of assortment and p is the proportion of the population that is Patrol. So when $r = 1$, the above evaluates to $\Pr(\text{Patrol}|\text{Patrol}) = 1$. Perfect assortment. But for $r < 1$, the amount of assortment depends upon how common the strategy is in the population. The corresponding chance Patrol meets Rest is:

$$\Pr(\text{Patrol}|\text{Rest}) = (1-r)(1-p)$$

Patrol meets rest when assortment fails, so no r term, and then meets Rest with chance $1-p$. These two probabilities are the statistical assortment model. They let us explore the general implications of assortment, without worrying yet about the mechanisms that produce it. That comes later.

So let's reconsider when Patrol can be evolutionarily stable. When common, the fitness of Patrol is $b - c/2$, because it meets itself, assortment or not. What is the fitness of a rare Rest? Rest will experience assortment, even when rare. It's fitness is:

$$\underbrace{(r + (1-r)(1-p))0}_{\text{Rest}|\text{Rest}} + \underbrace{(1-r)pb}_{\text{Rest}|\text{Patrol}}$$

When Patrol is common, $p \approx 1$, so the above simplifies to:

$$(1 - r)b$$

And so the condition for Patrol to be evolutionarily stable is:

$$b - c/2 > (1 - r)b$$

Let's rearrange this so it is a condition in terms of r:

$$r > \frac{c}{2b}$$

This can be satisfied, if r is large enough, even when $c > b$ (a prisoner's dilemma). For example, if $c = 3$ and $b = 2$, then this requires $r > 3/4$.

Assortment also changes the condition for Patrol to invade. When Rest is common, it always gets zero. A rare Patrol gets:

$$\underbrace{(r + (1 - r)p)(b - c/2)}_{\text{Patrol}|\text{Patrol}} + \underbrace{(1 - r)(1 - p)(b - c)}_{\text{Patrol}|\text{Rest}}$$

Now $p \approx 0$, so this simplifies to:

$$r(b - c/2) + (1 - r)(b - c)$$

For Patrol to invade a population, we require:

$$r(b - c/2) + (1 - r)(b - c) > 0$$

This condition simplifies to $r > 2(1 - b/c)$. This is not the same condition that Patrol needs for stability. But again, this can be satisfied when $c > b$. But it might require implausible amounts of assortment, so don't start celebrating.

Assortment behaves in different ways in different games. It changes stability conditions, as well as basins of attraction. Then details of the evolutionary process, such as the genetic architecture of the strategies, or whether strategies are learned rather than genetic, can influence whether cooperative strategies succeed in the long run.

Cooperation in large groups

[need a section on how cooperation in groups larger than pairs influences the different games]