

## Evolution, Cooperation, and Repeated Games (based on work with D. Fudenberg)

E. Maskin (Institute for Advanced Study)

The theory of repeated games has been an important tool in the behavioral and biological sciences. Indeed, it provides the central model for explaining how agents with selfish objectives might nevertheless behave cooperatively and efficiently in a long-term relationship. For that reason, it has been invoked often by economists, political scientists, ecologists, anthropologists and others interested in human cooperation.

Repeated-game theory offers a beautifully simple answer to the question of why selfish agents should cooperate: namely, they should do so to ensure continued cooperation in the future. For an illustration of this point, consider the classic embodiment of the tension between self-interest and cooperation, the Prisoner's Dilemma.

	<i>C</i>	<i>D</i>
<b>Cooperate</b>	2, 2	-1, 3
<b>Defect</b>	3, -1	<b>0, 0</b>

**Prisoner's Dilemma**

In this game there are two players—a row player and a column player—each of whom has two possible actions: to “cooperate” (*C*) or to “defect” (*D*). The payoffs to each combination of player's actions are given by the table above, where in each box the first number is the row player's payoff and the second is the column player's payoff. Notice that if the game is played just once, then regardless of what the other player does,

it is optimal (i.e., a *dominant strategy*) for each player to play *D*: by doing so he gets 3 rather than 2 if the other player plays *C*, and 0 rather than  $-1$  if the other player plays *D*. Thus, the predicted outcome of the one-shot game is for each player to play *D* (and thereby get a zero payoff), even though both would be better off if they played *C* (they would then each get 2).

Now imagine that the game is repeated many times (formally, infinitely many times), and that a player cares about his payoff on average in the long run rather than in any particular iteration. Unlike before, playing cooperatively may now be in the player's best interest. In fact, the strategy in which a player (i) plays *C* in the first iteration, (ii) continues to play *C* as long as both players have cooperated in the past, and (iii) plays *D* otherwise (call this strategy CC for "conditionally cooperate") generates an *equilibrium* in the sense that if both players adopt it, neither has the incentive to deviate to any other strategy unilaterally. To see this, note that when players both adhere to CC, a never-ending stream of cooperation results, giving each player an average payoff of 2 per period. If, however, a player unilaterally deviates to some other strategy, then either (a) in some iteration, he plays *D*—in which case his opponent always plays *D* thereafter, so that he (the deviator) obtains a zero average payoff at best, or else (b) despite the deviation, he continues to play *C* in every iteration—in which case, he still gets an average of 2, and so doesn't gain from deviating. Thus, repetition makes cooperation plausible as a possible outcome.

Unfortunately, cooperation is not the *only* plausible outcome. For example, the strategy in which a player *always plays D* (AD) also generates an equilibrium if both players adopt it. And these two extremes—CC, which leads to cooperation in every

iteration, and AD, which leads to no cooperation at all—aren't the only possibilities: the Folk Theorem of repeated games (c.f. Fudenberg and Maskin 1986) tells us that every intermediate possibility between full cooperation and full defection can occur in equilibrium as well. Moreover, the theory does not favor any particular equilibrium over the others. That is, the theory makes scarcely any clear-cut prediction about behavior at all. It is practically unfalsifiable.

Evolution (either biological or cultural) might be expected to help resolve this predictive murkiness. Evolutionary forces often foster efficiency, and so we could hope that uncooperative behavior will be rooted out and cooperative behavior promoted by selective pressure. Indeed, this view was set out forcefully in Axelrod and Hamilton (1981) and later in the popular book Axelrod (1984). The basic idea is easy to convey. Imagine that we start with a population in which all players use the strategy AD. Now, suppose that a small group of “mutants” are introduced who use the strategy CC. Notice that CC earns the same average payoff (namely, zero) against AD that AD does against itself. Moreover, CC earns an average payoff of 2 against itself, whereas AD gets only 0 on average against CC. Thus, in expectation (assuming that pairs of strategies are picked at random from the entire population), CC performs strictly better than AD. And so, according to standard evolutionary dynamics, CC will replicate faster than AD and eventually take over the population. That is, uncooperative behavior in the guise of AD will eventually be driven out.

Unfortunately, among uncooperative strategies, AD is unrepresentatively easy to drive out. Consider instead ALT, the strategy that alternates between *C* and *D* until someone breaks the alternating pattern, after which point it always plays *D*. Unlike AD,

a population of ALTs *cannot* be invaded by a mutant strategy. To see this, notice that a mutant would have to conform to the alternating pattern: otherwise, it would do strictly worse (an average payoff of 0) against ALT than ALT itself does against (average payoff 1), and so could not grow relative to ALT. But a mutant conforming to the alternating pattern would also fail to perform better in expectation than ALT. Thus, ALT is *evolutionarily stable*<sup>1</sup>(ES) despite being quite inefficient and uncooperative.<sup>2</sup>

Yet there is a sense in which ALT is too inflexible. It relies, after all, on perfect alternation: any deviation from the pattern  $C, D, C, D, C, \dots$  is “punished” by an infinite succession of  $D$ s. This suggests that it might not fare so well in an environment in which strategies are not always executed correctly, i.e., in which there is a small but positive probability that a strategy makes a mistake. Indeed, I claim that in such an environment, ALT is no longer ES.

Specifically, consider mutant strategy  $s'$  that is identical to ALT except when the alternating pattern has been broken. In the iteration just after the pattern is broken,  $s'$  plays  $C$  (unlike ALT, which plays  $D$ ) for one period - - to “signal” its willingness to cooperate. If the other player also plays  $C$  in that iteration, then  $s'$  plays  $C$  from then on. But if the other player plays  $D$ ,  $s'$  (like ALT) plays  $D$  thereafter.

I claim that  $s'$ , so constructed, will successfully invade a population of ALTs if strategies are subject to a small probability of inaccurate execution (so that, with positive probability, the alternating pattern will be broken). To see this, notice that (i)  $s'$  is

---

<sup>1</sup> Roughly speaking, a strategy  $s$  is ES if no mutant strategy  $s'$  performs better than  $s$  in expectation against a population consisting mostly of  $s$  but with a small proportion of  $s'$ .

<sup>2</sup> In fact, the situation is even worse than ALT would suggest. Consider the strategy that repeatedly follows the pattern  $C$  followed by *two*  $D$ s until the pattern is broken, at which point it thereafter plays  $D$ . For the same reason as ALT, this more elaborate strategy is ES, yet it attains an average payoff of only  $\frac{2}{3}$ . In fact, by continuing to add more  $D$ s to the repeated pattern, we can obtain an ES strategy that is arbitrarily close in average payoff to the fully uncooperative strategy AD.

identical to ALT before the alternating pattern is broken; (ii)  $s'$  and ALT each get a payoff of 0 per period against ALT after the pattern is broken, (iii)  $s'$  gets a payoff of almost 2 but ALT gets only 0 against  $s'$  after the pattern is broken. Thus, in expectation,  $s'$  performs better than ALT, and so will successfully invade.

In Fudenberg and Maskin (1990), (2007), Drew Fudenberg and I characterize the payoffs corresponding to evolutionarily stable strategies in two-player symmetric<sup>3</sup> repeated games when (i) there is a positive probability  $p$  in each iteration that a strategy is misexecuted, and (ii) players discount future payoffs at a positive rate  $r$ , so that instead of maximizing the long-run average payoff, a player maximizes the *discounted* average payoff

$$\frac{r}{1+r} \sum_{t=1}^{\infty} \left( \frac{1}{1+r} \right)^{t-1} \pi_t,$$

where  $\pi_t$  is his payoff in iteration  $t$ . So that I can state these characterization results reasonably precisely, let us define a pair of payoffs  $(v_R, v_C)$  (where  $v_R$  and  $v_C$  are the row and column players' payoffs respectively) in a two-player symmetric game  $g$  to be *strongly efficient* if (a)  $(v_R, v_C)$  are feasible payoffs for  $g$  and (b)  $v_R + v_C$  maximizes the sum of the two players' payoffs among all feasible payoffs for  $g$ . Thus in the Prisoner's Dilemma (see the table on p.1) the unique strongly efficient payoffs are (2, 2). But in the Battle of the Sexes, as given in the following table,

---

<sup>3</sup> A two-player game is symmetric if the two players have the same set of actions and if interchanging actions between the players causes the corresponding payoffs to be interchanged.

	<b>Ballet</b>	<b>Boxing</b>
<b>Boxing</b>	<b>0,0</b>	<b>1,2</b>
<b>Ballet</b>	<b>2,1</b>	<b>0,0</b>

**Battle of the Sexes**

any convex combination of (1,2) and (2,1) is strongly efficient (assuming that players can randomize between the two points). Let  $\underline{v}$  be the smallest payoff consistent with strong efficiency, i.e.,

$$\underline{v} = \min \{v_R \mid (v_R, v_C) \text{ is strongly efficient}\}.$$

Thus,  $\underline{v} = 2$  in the Prisoner's Dilemma, but  $\underline{v} = 1$  in the Battle of the Sexes.

The first result asserts that if the mistake probability and discount rate are small (but positive), the payoff resulting from an evolutionarily stable strategy cannot be much less than  $\underline{v}$ . More formally we have

*Theorem 1:* Given  $\varepsilon > 0$ , then if  $p > 0$  and  $r > 0$  are sufficiently small, the payoffs generated when both player use an ES strategy  $s$  cannot be less than  $\underline{v} - \varepsilon$ .

Note that Theorem 1 implies that, for the Prisoner's Dilemma, evolutionarily stable behavior must be almost fully cooperative. This is not the case, however, for the Battle of the Sexes. Full cooperation would entail playing (Ballet, Ballet) or (Boxing, Boxing), so that, on average, players would be getting a payoff of  $1\frac{1}{2}$  each. Evolutionary stability, by contrast, ensures only that average payoffs will not be (much) below 1.

The proof of Theorem 1 is similar to the argument above showing that ALT is not ES when  $p$  is positive. If  $s$  fails to attain the payoff  $\underline{v}$ , then we can construct a mutant

strategy  $s'$  that mimics  $s$  most of the time but exploits  $s$ 's inefficiency by “singling” its willingness to play more efficiently. If the other player reciprocates this signal, then  $s'$  will thereafter play strongly efficiently. If not, then  $s'$  will revert to playing like  $s$ .

Theorem 1, by itself, is an incomplete characterization because it doesn't deal with the issue of whether ES strategies actually exist. Accordingly, the second main result establishes that if  $v$  exceeds  $\underline{v}$ , there exists an ES strategy  $s$  attaining the payoff  $v$  (approximately), provided that  $r$  and  $p$  are small. More formally, we can state

*Theorem 2:* Let  $(v, v)$  be feasible payoffs with  $v \geq \underline{v}$ . For all  $\varepsilon > 0$ , there exist  $r$  and  $p$  sufficiently small so that there exists an ES  $s$  for which, if both players use  $s$ , the corresponding payoffs are within  $\varepsilon$  of  $(v, v)$ .

For the Prisoner's Dilemma, Theorem 2 asserts the existence of an ES strategy that (approximately) attains full cooperation. What form might such a strategy take? Like ALT, CC is too inflexible when  $p$  is positive: one mistake leads to  $D$  forever. The strategy Tit-for-tat (play  $C$  in the first iteration and thereafter do whatever the other player did in the previous iteration), emphasized by Axelrod and Hamilton (1981), is similarly too readily thrown off track by mistakes. If say, the row player (by mistake) plays  $D$  in the first iteration, Tit-for-tat will have the column player follow with  $D$  in the second iteration, and then will induce the row player to play  $D$  again in the third iteration, etc. That is, there will be a stream of  $D$ s that ends only when someone makes another mistake. A more robust strategy against mistakes takes the form: play  $C$  in the first iteration, and thereafter play  $C$  if *both* players played  $C$  the previous iteration or if *neither* did. Indeed, this strategy is ES for the payoffs given the Prisoner's Dilemma table above (see Fudenberg and Maskin 1990).

To see how ES strategies can generate payoffs quite short of full cooperation in games resembling the Battle of the Sexes, consider the game given by the following table

	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>
<i>a</i>	0,0	4,1	0,0	0,0
<i>b</i>	1,4	0,0	0,0	0,0
<i>c</i>	0,0	0,0	0,0	0,0
<i>d</i>	0,0	0,0	0,0	2,2

**Expanded Battle of the Sexes**

Consider the strategy  $s$  that (i) plays  $d$  as long as in every past iteration either both players played  $d$  or neither  $d$ , (ii) plays  $a$  forever if the other player was the first to deviate from  $d$ , and (iii) plays  $b$  forever if it itself was the first to deviate from  $d$ . It can be shown that  $s$  is ES for  $r$  and  $p$  sufficiently small. Nevertheless, the combination  $(d, d)$  generates payoffs  $(2, 2)$  that are not strongly efficient (strongly efficient payoffs sum to 5 in this example). The reason why, despite this inefficiency,  $s$  is invulnerable to invasion by a mutant is that the “punishment” for deviating from  $d$  is itself strongly efficient—i.e., play generates payoffs to  $(1,4)$  or  $(4,1)$ —and so there is no way that a mutant can get a toehold against  $s$ .

Bibliographical note: Theorems 1 and 2 are established by Fudenberg and Maskin (1990) for the case in which  $r$  and  $p$  are *infinitesimal*. Fudenberg and Maskin (2007) treat the case of finite values of  $r$  and  $p$ .

#### References

Axelrod, R. (1984), *The Evolution of Cooperation*, Basic Books.



Axelrod, R. and W. Hamilton (1981), "The Evolution of Cooperation," *Science*, 211: 1390-96.

Fudenberg, D. and E. Maskin (1986), "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica*, 54(3), pp.533-554. Reprinted in A. Rubinstein (ed.), *Game Theory in Economics*, London: Edward Elgar, 1995.

Fudenberg, D. and E. Maskin (1990), "Evolution and Cooperation in Noisy Repeated Games," *American Economic Review*, 80, 274-279.

Fudenberg, D. and E. Maskin (2007), "Evolution and Repeated Games," mimeo.