Prisoner's Dilemma

First published Thu Sep 4, 1997; substantive revision Mon Oct 22, 2007

Tanya and Cinque have been arrested for robbing the Hibernia Savings Bank and placed in separate isolation cells. Both care much more about their personal freedom than about the welfare of their accomplice. A clever prosecutor makes the following offer to each. "You may choose to confess or remain silent. If you confess and your accomplice remains silent I will drop all charges against you and use your testimony to ensure that your accomplice does serious time. Likewise, if your accomplice confesses while you remain silent, they will go free while you do the time. If you both confess I get two convictions, but I'll see to it that you both get early parole. If you both remain silent, I'll have to settle for token sentences on firearms possession charges. If you wish to confess, you must leave a note with the jailer before my return tomorrow morning."

The "dilemma" faced by the prisoners here is that, whatever the other does, each is better off confessing than remaining silent. But the outcome obtained when both confess is worse for each than the outcome they would have obtained had both remained silent. A common view is that the puzzle illustrates a conflict between individual and group rationality. A group whose members pursue rational self-interest may all end up worse off than a group whose members act contrary to rational self-interest. More generally, if the payoffs are not assumed to represent self-interest, a group whose members rationally pursue any goals may all meet less success than if they had not rationally pursued their goals individually. A closely related view is that the prisoner's dilemma game and its multi-player generalizations model familiar situations in which it is difficult to get rational, selfish agents to cooperate for their common good. Much of the contemporary literature has focused on identifying conditions under which players would or should make the "cooperative" move corresponding to remaining silent. A slightly different interpretation takes the game to represent a choice between selfish behavior and socially desirable altruism. The move corresponding to confession benefits the actor, no matter what the other does, while the move corresponding to silence benefits the other player no matter what that player does. Benefiting oneself is not always wrong, of course, and benefiting others at the expense of oneself is not always morally required, but in the prisoner's dilemma game both players prefer the outcome with the altruistic moves to that with the

selfish moves. This observation has led David Gauthier and others to take the Prisoner's Dilemma to say something important about the nature of morality.

Puzzles with the structure of the prisoner's dilemma were devised and discussed by Merrill Flood and Melvin Dresher in 1950, as part of the Rand Corporation's investigations into game theory (which Rand pursued because of possible applications to global nuclear strategy). The title "prisoner's dilemma" and the version with prison sentences as payoffs are due to Albert Tucker, who wanted to make Flood and Dresher's ideas more accessible to an audience of Stanford psychologists. Although Flood and Dresher didn't themselves rush to publicize their ideas in external journal articles, the puzzle attracted widespread attention in a variety of disciplines. Christian Donninger reports that "more than a thousand articles" about it were published in the sixties and seventies. A bibliography (Axelrod and D'Ambrosio) of writings between 1988 and 1994 that pertain to Robert Axelrod's research on the subject lists 209 entries. Since then the flow has shown no signs of abating.

The sections below provide a variety of more precise characterizations of the prisoner's dilemma, beginning with the narrowest and survey some connections with similar games and some applications in philosophy and elsewhere. Particular attention is paid to "evolutionary" versions of the game in which members of a population play one another repeatedly and those who get higher payoffs "reproduce" more rapidly than those who get lower payoffs. 'Prisoner's dilemma' is abbreviated as 'PD'.

- 1. Symmetric 2×2 PD With Ordinal Payoffs
- 2. Asymmetry
- 3. Multiple Moves
- 4. Multiple Players
- 5. Single Person Interpretations
- 6. Cardinal Payoffs
- 7. The PD with Replicas and Causal Decision Theory
- 8. The Stag Hunt and the PD
- 9. Asynchronous Moves
- 10. Transparency
- 11. Finite Iteration
- 12. The Centipede and the Finite IPD
- 13. Infinite Iteration
- 14. Indefinite Iteration
- 15. Iteration With Error
- 16. Evolution
- 17. Spatial PDs
- 18. PDs and Social Networks
- 19. Group Selection and the Haystack PD

1. Symmetric 2×2 PD With Ordinal Payoffs

In its simplest form the PD is a game described by the payoff matrix:

	C	D
C	R, R	S, T
D	T, S	P, P

satisfying the following chain of inequalities:

    PD1) $T>R>P>S$

There are two players, Row and Column. Each has two possible moves, "cooperate" or "defect," corresponding, respectively, to the options of remaining silent or confessing in the illustrative anecdote above. For each possible pair of moves, the payoffs to Row and Column (in that order) are listed in the appropriate cell. R is the "reward" payoff that each player receives if both cooperate. P is the "punishment" that each receives if both defect. T is the "temptation" that each receives if he alone defects and S is the "sucker" payoff that he receives if he alone cooperates. We assume here that the game is symmetric, i.e., that the reward, punishment, temptation or sucker payoff is the same for each player, and payoffs have only ordinal significance, i.e., they indicate whether one payoff is better than another, but tell us nothing about how much better. It is now easy to see that we have the structure of a dilemma like the one in the story. Suppose Column cooperates. Then Row gets R for cooperating and T for defecting, and so is better off defecting. Suppose Column defects. Then Row gets S for cooperating and P for defecting, and so is again better off defecting. The move **D** for Row is said to *strictly dominate* the move **C**: whatever his opponent does, he is better off choosing **D** than **C**. By symmetry **D** also strictly dominates **C** for Column. Thus two "rational" players will defect and receive a payoff of P, while two "irrational" players can cooperate and receive greater payoff R. In standard treatments, game theory assumes rationality and common knowledge. Each player is rational, knows the other is rational, knows that the other knows he is rational, etc. Each player also knows how the other values the outcomes. But since **D** strictly dominates **C** for both players, the argument for dilemma here requires only that each player knows his own payoffs. (The argument remains valid, of course, under the stronger standard assumptions.) It is also

worth noting that the outcome ($\mathbf{D}$, $\mathbf{D}$) of both players defecting is the game's only strong nash equilibrium, i.e., it is the only outcome from which each player could only do worse by unilaterally changing its move. Flood and Dresher's interest in their dilemma seems to have stemmed from their view that it provided a counterexample to the claim that the nash equilibria of a game constitute its natural "solutions".

If there can be "ties" in rankings of the payoffs, condition PD1 can be weakened without destroying the nature of the dilemma. For suppose that one of the following conditions obtains:

> PD2) $T>R>P\geq S$, or
> $T\geq R>P>S$

Then, for each player, although $\mathbf{D}$ does not strictly dominate $\mathbf{C}$, it still *weakly dominates* in the sense that each player always does at least as well, and sometimes better, by playing $\mathbf{C}$. Under these conditions it still seems rational to play $\mathbf{D}$, which again results in the payoff that neither player prefers. Let us call a game that meets PD2 a *weak PD*. Note that in a weak PD that does not satisfy PD1 mutual defection is no longer a nash equilibrium in the strong sense defined above. It is still, however, the only nash equilibrium in the weaker sense, that neither player can improve its position by unilaterally changing its move. Again, one might suppose that if there is a unique nash equilibrium of this weaker variety, rational self-interested players would reach it.

2. Asymmetry

Without assuming symmetry, the PD can be represented by using subscripts r and c for the payoffs to Row and Column.

	C	D
C	R_r, R_c	S_r, T_c
D	T_r, S_c	P_r, P_c

If we assume that the payoffs are ordered as before for each player, i.e., that $T_i>R_i>P_i>S_i$ when $i=r, c$, then, as before, $\mathbf{D}$ is the strictly dominant move for both players, but the outcome ($\mathbf{D}$, $\mathbf{D}$) of both players making this move is worse for each than ($\mathbf{C}$, $\mathbf{C}$). The force of the dilemma can now also be felt under weaker conditions, however. Consider the following three pairs of inequalities:

> PD3) a. $T_r>R_r$ and $P_r>S_r$

$\quad$ b. $T_c > R_c$ and $P_c > S_c$
$\quad$ c. $R_r > P_r$ and $R_c > P_c$

If these conditions all obtain the argument for dilemma goes through as before. Defection strictly dominates cooperation for each player, and (C, C) is strictly preferred by each to (D, D). If one of the two > signs in each of the conditions a - c is replaced by a weak inequality sign ≥ we have a weak PD. D weakly dominates C for each player (i.e., D is as good as C in all cases and better in some) and (C, C) weakly better than (D, D) (i.e., it is at least as good for both players and better for one). Since none of the clauses requires comparisons between r's payoffs and c's, we need not assume that > has any "interpersonal" significance.

Now suppose we drop the first inequality of either a or b (but not both). A game that meets the resulting conditions might be termed a *common knowledge PD*. As long as each player knows that the other is rational and each knows the other's ordering of payoffs, we still feel the force of the dilemma. For suppose a holds. Then D is the dominant move for Row. Column, knowing that Row is rational, knows that Row will defect, and so, by the remaining inequality in b, will defect himself. Similarly, if b holds Column will defect, and Row, realizing this, will defect herself. By c, the resulting (D, D) is again worse for both than (C, C).

3. Multiple Moves

Speaking generally, one might say that a PD is a game in which a "cooperative" outcome obtainable only when every player violates rational self-interest is unanimously preferred to the "selfish" outcome obtained when every player adheres to rational self-interest. We can characterize the selfish outcome either as the result of each player pursuing its dominant (strongly dominant) strategy, or as the unique weak (strong) nash equilibrium. In a two move game the two characterizations come to the same thing — a dominant move pair is a unique equilibrium and a unique equilibrium is a dominant move pair. As the payoff matrix below shows, however, the two notions diverge in a game with more than two moves.

	C	D	N
C	R, R	S, T	T, S
D	T, S	P, P	R, S
N	S, T	S, R	S, S

Here each player can choose "cooperate", "defect" or "neither" and the payoffs are ordered as before. Defection is no longer dominant, because each player is better off choosing **C** than **D** when the other chooses *N*. Nevertheless (**D**, **D**) is still the unique equilibrium. Let us label a game like this in which the selfish outcome is the unique equilibrium an *equilibrium PD*, and one in which the selfish outcome is a pair of dominant moves a *dominance PD*. As will be seen below, attempts to "solve" the PD by allowing conditional strategies can create multiple-move games that are themselves equilibrium PDs.

4. Multiple Players

Most of those who maintain that the PD illustrates something important about morality seem to believe that the basic structure of the game is reflected in situations that larger groups, perhaps entire societies, face. The most obvious generalization from the two-player to the many-player game would pay each player *R* if all cooperate, *P* if all defect, and, if some cooperate and some defect, it would pay the cooperators *S* and the defectors *T*. But it is unlikely that we face many situations of this structure.

A common view is that a multi-player PD structure is reflected in what Garret Hardin popularized as "the tragedy of the commons." Each member of a group of neighboring farmers prefers to allow his cow to graze on the commons, rather than keeping it on his own inadequate land, but the commons will be rendered unsuitable for grazing if it is used by more than some threshold number use it. More generally, there is some social benefit *B* that each member can achieve if sufficiently many pay a cost *C*. We might represent the payoff matrix as follows:

	more than *n* choose C	*n* or fewer choose C
C	*C+B*	*C*
D	*B*	0

The cost *C* is assumed to be a negative number. The "temptation" here is to get the benefit without the cost, the reward is the benefit with the cost, the punishment is to get neither and the sucker payoff is to pay the cost without realizing the benefit. So the payoffs are ordered $B > (B+C) > 0 > C$. As in the two-player game, it appears that **D** strongly dominates **C** for all players, and so rational players would choose **D** and achieve 0, while preferring that everyone would choose **C** and obtain *C+B*.

Unlike the more straightforward generalization, this matrix does reflect common social

choices — between depleting and conserving a scarce resource, between using polluting and non-polluting means of manufacture or disposal, and between participating and not participating in a group effort towards some common goal. When *n* is small, it represents a version of what has been called the "volunteer dilemma". A group needs a few volunteers, but each member is better off if others volunteer. (Notice, however, that in a true volunteer dilemma, where only one volunteer is needed, *n* is zero and the top right outcome is impossible. Under these conditions **D** no longer dominates **C** and the game loses its PD flavor.) A particularly vexing manifestation of this game occurs when a vaccination known to have serious risks is needed to prevent the outbreak of a fatal disease. If enough of her neighbors get the vaccine, each person may be protected without assuming the risks.

The tragedy of the commons game diagramed above has a somewhat different character than the two-player PD. First, even if each player's moves are entirely independent of the others, the alternatives represented by the columns in the commons matrix above are no longer independent of the alternatives represented by the rows. My choosing **C** necessarily increases the chances that more than *n* people will choose **C**. To ensure independence we should really redraw the matrix as follows:

	more than *n* others choose **C**	*n* others choose **C**	fewer than *n* others choose **C**
C	*C+B*	*C+B*	*C*
D	*B*	0	0

But now we see that move **D** does *not* dominate **C**. When we are at the threshold of adequate cooperation, where exactly *n* others choose **C**, I am better off cooperating. Provided that *n* is large, however, it would seem that this effect could be ignored and we could assume, for practical purposes, that the payoff matrix is like the previous one.

Similarly, whereas we saw in the original PD that mutual defection was the only nash equilibrium, this game has two equilibria. One is universal defection, since any player unilaterally departing from that outcome will move from payoff 0 to *C*. But a second is the state of *minimally effective* cooperation, where the number of cooperators just exceeds the threshold. A defector who unilaterally departs from that outcome will move from *B* to *B+C* and a cooperator who unilaterally departs will move from *B+C* to 0. This might suggest that the tragedy of the commons is less tragic than the PD, but in real life situations, it would seem unlikely that the participants would know when they are at the equilibrium point of minimally effective cooperation.

Furthermore, in the ordinary PD, universal cooperation is a pareto-optimal outcome, i.e.,

there is no outcome in which each player is at least as well off and some are better off. But in the commons game the only pareto optimal outcomes are those of minimally effective cooperation. Whether universal cooperation is nevertheless desirable may depend on the nature of the choices involved. In the medical example it may seem best to vaccinate everyone. In the agricultural example, however it seems foolish to stipulate that nobody use the commons. Someone who avoids vaccination in the former case is seen as a "free rider". An underused commons in the latter seems to exemplify "surplus cooperation."

The two-person version of the tragedy of the commons game (with threshold of one) produces a matrix presenting considerably less of a dilemma.

	C	D
C	$B+C, B+C$	$C, 0$
D	$0, C$	$0, 0$

This game captures David Hume's example of a boat with one oarsman on the port side and another on the starboard (provided we assume that Hume's oarsmen must make their choices between rest and exertion simultaneously). Mutual cooperation is identical to minimally effective cooperation and therefore is both an equilibrium outcome and a pareto optimal outcome. Games of this sort will be discussed below under the label "Stag Hunt."

The above representations of the tragedy of the commons make the simplifying assumptions that the costs and benefits of cooperation are the same for each player, that the cost of cooperation is independent of the number of players who cooperate, and that the size of the benefit (0 or B) depends only on whether the number of cooperators exceeds the threshold. A somewhat more general account would replace C and B by functions $C(i, j)$ and $B(i, j)$ representing the cost of cooperation to player i when he is one of exactly j players who cooperate and the benefit to player i when exactly j players cooperate. We suppose that there is some threshold t for minimally effective cooperation so that $B(i, j)$ is not defined unless $j > t$. We may also assume additional cooperation never reduces the benefit i gets from effective cooperation, i.e., $B(i, j+1) \geq B(i, j)$ when $j > t$ and that additional defection never reduces the cost i bears in cooperating, i.e., $C(i, j+1) \geq C(i, j)$. Now suppose, in addition, that, once the threshold of effective cooperation has been exceeded, any benefit one gets from from the presence of an additional cooperator is exceeded by one's cost of cooperation and that the costs of ineffective cooperation are genuine, i.e., for all players i, $B(i, j) > (B(i, j+1) + C(i, j+1))$ when j is greater than t and $0 > C(i, j)$ when j is less than or equal to t. Finally, suppose that the benefits to each player i, of effective cooperation exceed the costs, i.e., for $j > t$, $B(i, j) + C(i, j) > 0$ We then have a tragedy of the commons

game, which presents a familiar dilemma: defection benefits an individual in every circumstance (except the one where exactly t others cooperate) but everybody is better off in any state of effective cooperation than in any state without it. This account could be easily be modified to allow threshold of minimally effective cooperation to differ from one individual to another (i's clean water requirements might be more stringent than j's for example) or to allow B to be defined everywhere (thus eliminating the threshold, so that we always benefit from another's cooperation). The resulting game would still have its PD flavor.

Phillip Pettit has pointed out that examples that might be represented as many-player PDs come in two flavors. The examples discussed above might be classified as free-rider problems. My temptation is to enjoy some benefits brought about by burdens shouldered by others. The other flavor is what Pettit calls "foul dealer" problems. My temptation is to benefit myself by hurting others. Suppose, for example, that a group of people are applying for a single job, for which they are equally qualified. If all fill out their applications honestly, they all have an equal chance of being hired. If one lies, however, he can ensure that he is hired while, let us say, incurring a small risk of being exposed later. If everyone lies, they again have an equal chance for the job, but now they all incur the risk of exposure. Thus a lone liar, by reducing the others' chances of employment from slim to none, raises his own chances from slim to sure. As Pettit points out, when the minimally effective level of cooperation is the same as the size of the population, there is no opportunity for free-riding (everyone's cooperation is needed), and so the PD must be of the foul-dealing variety. But (Pettit's contrary claim notwithstanding) not all foul-dealing PDs seem to have this feature. Suppose, for example, that two applicants in the story above will be hired. Then everyone gets the benefit (a chance of employment without risk of exposure) unless two or more players lie. Nevertheless, the liars seem to be foul dealers rather than free riders. A better characterization of the foul-dealing dilemma might be that every defection from a generally cooperative state strictly reduces the payoffs to the cooperators, i.e., for every player i and every j greater than the threshold, either $B(i,j+1)+C(i,j+1)>B(i,j)+C(i,j)$. A free-rider's defection benefits himself but does not, by itself, hurt the cooperators. A foul-dealer's defection benefits himself and hurts the cooperators.

The game labeled a many-person PD in Schelling, Molander 1992 and elsewhere requires that the payoff to each co-operator and defector increases strictly with the number of cooperators and that the sum of the payoffs to all parties increases with the number of cooperators (so that one party's switching from defection to cooperation always raises the sum). Neither of these conditions is met by the formulation and the examples discussed above. They may, however, hold "locally," i.e., for j close to the threshold t for minimally effective cooperation, it may be reasonable to assume that:

- for every individual i, $B(i,j+1)+C(i,j+1)>B(i,j)+C(i,j)$ for $j>t$,

- for every individual i, $C(i,j+1) > C(i,j)$ for $j \leq t$, and
- $B(1,j+1) + C(1,j+1) + \ldots + B(j+1,j+1) + C(t+1,j+1) + B(j+2,j+1)) + \ldots + B(n,j+1) > B(1,j) + C(1,j) + \ldots + B(j,j) + C(j,j) + B(j+1,j) + \ldots + B(n,j)$.

By requiring that cooperation of others always benefits each and all, the Schelling and Molander formulations of the *n*-person PD fail to model the surplus cooperation/free rider phenomenon that seems to infuse "Tragedy of the Commons" situations. Their conditions might, however be a plausible model for certain *public good* dilemmas. It is not unreasonable to suppose that *any* contribution towards public health, national defense, highway safety, or clean air is valuable to everybody, no matter how little or how much we already have, but that the cost to each individual for his own contribution to those goods always exceeds the benefit that he derives from that contribution. This outlook has the advantage of focusing attention on the PD quality of the game. Defection dominates cooperation, while universal cooperation is unanimously preferred to universal defection. Michael Taylor goes even further in this direction. His version of the many-person PD requires only the two PD-conditions just mentioned and the one additional condition that defectors are always better off when some cooperate than when none do. (Taylor's main concern is with the iterated version of this game, a topic that will be addressed in future editions of this entry.)

5. Single Person Interpretations

The PD is usually thought to illustrate conflict between individual and collective rationality, but the multiple player form (or something very similar) has also been interpreted as demonstrating problems within standard conceptions of individual rationality. One such interpretation, elucidated in Quinn, derives from an example of Parfit's. A medical device enables electric current to be applied to a patient's body in increments so tiny that there is no perceivable difference between adjacent settings. You are attached to the device and given the following choice every day for ten years: advance the device one setting and collect a thousand dollars, or leave it where it is and get nothing. Since there is no perceivable difference between adjacent settings, it is apparently rational to advance the setting each day. But at the end of ten years the pain is so great that a rational person would sacrifice all his wealth to return to the first setting.

We can view the situation here as a multi-player PD in which each "player" is the temporal stage of a single person. So viewed, it has at least two features that were not discussed in connection with the multi-player examples. First, the moves of the players are sequential rather than simultaneous (and each player has knowledge of preceding moves). Second, there is the matter of gradation. Increases in electric current between adjacent settings are imperceptible, and therefore irrelevant to rational decision-making, but sums of a number such increases are noticeable and highly relevant. Neither of these features, however, is

peculiar to one-person examples. Consider, for example, the choice between a polluting and non-polluting means of waste disposal. Each resident of a lakeside community may dump his or her garbage in the lake or use a less convenient landfill. It is reasonable to suppose that each acts in the knowledge of how others have acted before. (See "Asynchronous Moves" below.) It is also reasonable to suppose that addition of one can of garbage to the lake has no perceptible effect on water quality, and therefore no effect on the welfare of the residents. The fact that the dilemma remains suggests that PD-like situations sometimes involve something more than a conflict between individual and collective rationality. In the one-person example, our understanding that we care more about our overall well-being than that of our temporal stages does not (by itself) eliminate the argument that it is rational to continue to adjust the setting. Similarly, in the pollution example, a decision to let collective rationality override individual rationality may not eliminate the argument for excessive dumping. It seems appropriate, however, to separate this issue from that raised in the standard PD. Gradations that are imperceptible individually, but weighty en masse give rise to intransitive preferences. This is a challenge to standard accounts of rationality whether or not it arises in a PD-like setting.

A second one-person interpretation of the PD is suggested in Kavka, 1991. On Kavka's interpretation, the prisoners are not temporal stages, but rather "subagents" reflecting different desiderata that I might bring to bear on a decision. Let us imagine that I am hungry and considering buying a snack. The options open to me are:

a. Buy a scoop of chocolate gelato.
b. Buy a scoop of orange sherbet.
c. Buy a granola bar.
d. Buy nothing.

My health-conscious side, "Arnold," orders these options in the following order: c, b, d, a. My taste-conscious side, "Eppie," ranks them: a, b, d, c. Such inner conflict among preferences might often be resolved in ways consistent with standard views about individual choice. My overall preference ordering, for example, might be determined from a weighted average of the utilities that Arnold and Eppie assign to each of the options. It is also possible, Kavka suggests, that my inner conflicts are resolved as if they were a result of strategic interaction among rational subagents. In this case, Arnold and Eppie can each choose either to *insist* on getting their way (**I**) or to *acquiesce* to a compromise (**A**). The interaction between subagents can then be represented by the following payoff matrix, where Arnold plays row and Eppie plays column.

	A	**I**
A	b	a

I	*c*	*d*

Examination of the table and preference orderings confirms that we again have an intrapersonal PD. Kavka argues that a story like this might "provide a psychologically plausible picture of how internal conflict can lead to suboptimal action." It also undermines a standard view that choices reflect values in favor of one that they partially reflect, "the structure of inner conflict".

6. Cardinal Payoffs

If the game specifies absolute (as opposed to relative) payoffs, then universal cooperation may not be a pareto optimal outcome even in the two person PD. For under some conditions both players do better by adopting a *mixed* strategy of cooperating with probability *p* and defecting with probability (1-*p*). This point is illustrated in the graphs below.
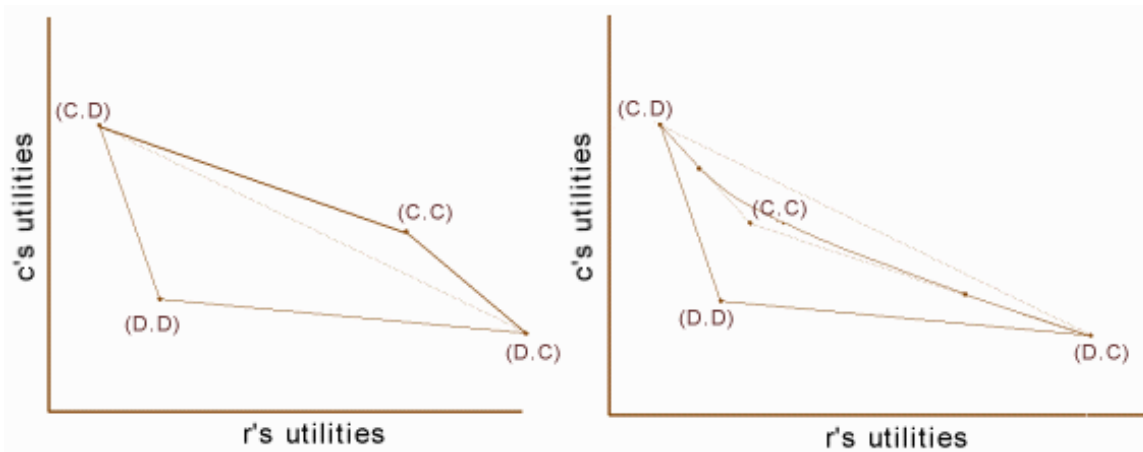

Figure 1

Here the *x* and *y* axes represent the utilities of Row and Column. The four outcomes entered in the matrix of the second section are represented by the labeled dots. Conditions PD3a and PD3b ensure that (**C**, **D**) and (**D**, **C**) lie northwest and southeast of (**D**, **D**), and PD3c is reflected in the fact that (**C**, **C**) lies northeast of (**D**, **D**). Suppose first that (**D**, **D**) and (**C**, **C**) lie on opposite sides of the line between (**C**, **D**) and (**D**, **C**), as in the graph on the left. Then the four points form a convex quadrilateral, and the payoffs of the feasible outcomes of mixed strategies are represented by all the points on or within this quadrilateral. Of course a player can really only get one of four possible payoffs each time the game is played, but the points in the quadrilateral represent the *expected values* of the payoffs to the two players. If Row and Column cooperate with probabilities *p* and *q* (and defect with probabilities *p*=1-p* and *q*=1-q*), for example, then the expected value of the payoff to Row is *p*qT+pqR+p*q*P+pq*S*. A rational self-interested player, according to a

standard view, should prefer a higher expected payoff to a lower one. In the graph on the left the payoff for universal cooperation (with probability one) is pareto optimal among the payoffs for all mixed strategies. In the graph on the right, however, where both $(\mathbf{D}, \mathbf{D})$ and $(\mathbf{C}, \mathbf{C})$ lie southwest of the line between $(\mathbf{C}, \mathbf{D})$ and $(\mathbf{D}, \mathbf{C})$, the story is more complicated. Here the payoffs of the feasible outcome lie within a figure bounded on the northeast by three distinct curve segments, two linear and one concave. Notice that $(\mathbf{C}, \mathbf{C})$ is now in the interior of the region bounded by solid lines, indicating that there are mixed strategies that provide both players a higher expected payoff than $(\mathbf{C}, \mathbf{C})$. It is important to note that we are talking about independent mixed strategies here. Row and Column use private randomizing devices and have no communication. If they were able to correlate their mixed strategies, so as to ensure, say $(\mathbf{C}, \mathbf{D})$ with probability p and $(\mathbf{D}, \mathbf{C})$ with probability p^*, the set of feasible solutions would extend up to (and include) the dotted line between $(\mathbf{C}, \mathbf{D})$ and $(\mathbf{D}, \mathbf{C})$. The point here is that, even confined to independent strategies, there are some games satisfying PD3 in which both players can both do better than they do with universal cooperation. A PD in which universal cooperation is pareto optimal may be called a pure PD. (This phenomenon is identified in Kuhn and Moresi and applied to moral philosophy in Kuhn 1996.) A pure PD is characterized by adding to PD3 the following condition.

P) $(T_r\text{-}R_r)(T_c\text{-}R_c)\leq(R_r\text{-}S\ _r)(R_c\text{-}S_c)$

In a symmetric game P reduces to the simpler condition

RCA) $R\geq1/2(T+S)$

(named after the authors Rapoport, Chammah and Axelrod who employed it).

7. The PD with Replicas and Causal Decision Theory

One controversial argument that it is rational to cooperate in a PD relies on the observation that my partner in crime is likely to think and act very much like I do. (See, for example, Davis 1977 and 1985 for a sympathetic presentation of one such argument and Binmore 1994,Chapters 3.4 and 3.5, for a reformulation and extended rebuttal.) In the extreme case, my accomplice is an exact *replica* of me who is wired just as I am so that, of necessity, we do the same thing. It would then seem that the only two possible outcomes are where both players cooperate and where both players defect. Since the reward payoff exceeds the punishment payoff, I should cooperate. More generally, even if my accomplice is not a perfect replica, the odds of his cooperating are greater if I cooperate and the odds of his defecting are greater if I defect. When the correlation between our behaviors is sufficiently strong or the differences in payoffs is sufficiently great, my *expected* payoff (as that term is usually understood) is higher if I cooperate than if I defect. The counter argument, of

course, is that my action is *causally independent* of my replica's. Since I can't affect what my accomplice does and since, whatever he does, my payoff is greater if I defect, I should defect. These arguments closely resemble the arguments for two positions on the *Newcomb Problem*, a puzzle popularized among philosophers in Nozick. (The extent of the resemblance is made apparent in Lewis.) The Newcomb Problem asks us to consider two boxes, one transparent and one opaque. In the transparent box we can see a thousand dollars. The opaque box may contain either a million dollars or nothing. We have two choices: take the contents of the opaque box or take the contents of both boxes. We know before choosing that a reliable predictor of our behavior has put a million dollars in the opaque box if he predicted we would take the first choice and left it empty if he predicted we would take the second. To see that each player in a PD faces a Newcomb problem, consider the following payoff matrix.

	C	D
C	m, m	$0, m+t$
D	$m+t, 0$	t, t

By "cooperating" (choosing the opaque box), each player ensures that the other gets a million dollars (and a thousand extra for defecting). By "defecting" (choosing both boxes) each player ensures that he will get thousand dollars himself (and a million more if the other cooperates). As long as $m>t>0$, the structure of this game is an ordinary two-player, two-move PD (and any such PD can be represented in this form). Furthermore, the arguments for "one-boxing" and "two-boxing" in a Newcomb problem are the same as the arguments for cooperating and defecting in a Prisoner's dilemma where there is positive correlation between the moves of the players. Two boxing is a *dominant* strategy: two boxes are better than one whether the first one is full or empty. On the other hand, if the predictor is reliable, the *expected* payoff for one-boxing is greater than the expected payoff for two-boxing. [See Hurley, however, for an argument that the two puzzles are significantly different. In a PD (of either the ordinary or the Newcombized variety) each player knows that the other is rational and that the other ranks the outcomes in the ways described. This, Hurley argues, opens a possibility for cooperative joint action that is absent in the original Newcomb problem.]

The intuition that two-boxing is the rational choice in a Newcomb problem, or that defection is the rational choice in the PD with positive correlation between the players' moves, seems to conflict with the idea that rationality requires maximizing expectation. This apparent conflict has led some to suggest that standard decision theory needs to be refined in cases in which an agent's actions provide *evidence* for, without *causing*, the

context in which he is acting. In the case of the PD, standard (evidential) decision theory asks Player One to compare his expected utilities of cooperation and defection, which can be written as $p(\mathbf{C}_2|\mathbf{C}_1) \times R + p(\mathbf{D}_2|\mathbf{C}_1) \times S$ and $p(\mathbf{C}_2|\mathbf{D}_1) \times T + p(\mathbf{D}_2|\mathbf{D}_1) \times P$ (where, for example, $p(\mathbf{C}_2|\mathbf{C}_1)$ is the conditional probability that player Two cooperates given that Player One cooperates). If the players' moves are strongly correlated then $p(\mathbf{C}_2|\mathbf{C}_1)$ and $p(\mathbf{D}_2|\mathbf{D}_1)$ will be close to one and $p(\mathbf{C}_2|\mathbf{D}_1)$ and $p(\mathbf{D}_2|\mathbf{C}_1)$ will be close to zero. On the suggested revision, these conditional probabilities should be replaced by some kind of causally conditional probabilities, which might (on some accounts) be expressed by phrases like "the probability that if One were to cooperate, Two would also cooperate." When the moves are causality independent this would just be the probability that Two cooperates.

The rather far-fetched scenario described in Newcomb's Problem initially led some to doubt the importance of the distinction between causal and evidential decision theory. Lewis argues that the link to the PD suggests that situations where the two decisions diverge are not so unusual, and recent writings on causal decision theory contain many examples far less bizarre than Newcomb's problem. (See Joyce, for example.)

It might be noted that what is here called "PD between replicas" is usually called "PD with twins" in the literature. One reason for the change in nomenclature is to distinguish these ideas from an experimental literature reporting on PD games played with real (identical or fraternal) twins. (See, for example, Segal and Hershberger.) It turns out that twins *are* more likely to cooperate in a PD than strangers, but there seems to be no suggestion that the reasoning that leads them to do so follows the controversial arguments presented above.

8. The Stag Hunt and the PD

The idea mentioned in the introduction that the PD models a problem of cooperation among rational agents is sometimes criticized because, in a true PD, the cooperative outcome is not a nash equilibrium. Any "problem" of this nature, the critics contend, would be an unsolvable one. (See for example, Sugden or Binmore 2005, Chapter 4.5.) By changing the payoff structure of the PD slightly, so that the reward payoff exceeds the temptation payoff, we obtain a game where mutual cooperation, as well as mutual defection, is a nash equilibrium. This game is known as the Stag Hunt. It might provide a better model for situations where cooperation is difficult, but still possible, and it may also be a better fit for other roles sometimes assigned to the PD. More specifically, a Stag Hunt is a two player, two move game with a payoff matrix like that for the PD pictured above where the conditions PD1 are replaced:

SH) a. $R > T$
        b. $R > P$

   c. *P>S*.

The fable dramatizing the game and providing its name, gleaned from a passage in Rousseau's *Discourse on Inequality*, concerns a hunting expedition rather than a jail cell interrogation. Two hunters are are looking to bag a stag. Success is uncertain and, if it comes, require the efforts of both. On the other hand, either hunter can forsake his partner and catch a hare with a good chance of success. A typical payoff matrix is shown below.

	C	D
C	4,4	0,3
D	3,0	3,3

Here the "cooperative" move is hunting stag with one's partner and "defection" is hunting hare by oneself. The "temptation" payoff in a Stag Hunt is no longer much of a temptation, but we retain the payoff terminology for ease of exposition. In this case the temptation and punishment penalties are identical, perhaps reflecting the fact that my partner's actions have no effect on my hunt for hare. Alternatively we could have temptation exceeding punishment, perhaps because hunting hare is more rewarding together than alone (though still less rewarding, of course, than hunting stag together), or we could have punishment exceeding temptation, perhaps because a second hare hunter represents unhelpful competition. Either way, the essence of the Stag Hunt remains. There are two equilibria, one unanimously preferred to the other. The Stag Hunt becomes a "dilemma" when rationality dictates that both players choose the action leading to the inferior equilibrium. It is clear that if I am certain that my partner will hunt stag I should join him and that if I am certain that he will hunt hare I should hunt hare as well. For this reason games with this structure are sometimes called games of "assurance" or "trust." If I do not know what my partner will do, standard decision theory tells me to maximize expectation. This requires, however, that I estimate the probability of my partner playing **C** or **D**. If I lack information to form any such estimates, then one putative principle of rationality ("indifference") suggests that I ought to treat all options as equally likely. By this criterion I ought to hunt hare if and only if the following condition is met:

   SHD) *T+P>R+S*

When SHD obtains, hare hunting is said to be the "risk-dominant" equilibrium. Let us call a Stag Hunt game where this condition is met a *Stag Hunt Dilemma*. The matrix above provides one example.

Another proposed principle of rationality ("maximin") suggests that I ought to consider the

worst payoff I could obtain under any course of action, and choose that action that maximizes this value. Since the sucker payoff is the worst payoff in a Stag Hunt, this principle suggests that *any* Stag Hunt presents a dilemma. Maximin, however, makes more sense as a principle of rationality for zero sum games, where it can be assumed that a rational opponent is trying to minimize my score, than for games like Stag Hunt, where a rational opponent may be quite happy to see me do well, as long as he does so as well.

The Stag Hunt can be generalized in the obvious way to accommodate asymmetric and cardinal payoffs. The quadrilateral formed by the games' graphical representation is convex, so the pure/impure distinction no longer applies. (I.e., in a Stag Hunt no mixed strategies are ever preferred to mutual cooperation.) The most obvious way to generalize the game to many players would retain the condition that there be exactly two equilibria, one unanimously preferred to the other. This might be a good model for cooperative activity in which success requires full cooperation. Imagine, for example, that a single polluter would spoil a lake, or a single leak would thwart an investigation. If many agents are involved and, by appeal to indifference or for other reasons, we estimate a fifty-fifty chance of cooperation from each, then these examples would represent Stag Hunt Dilemmas in an extreme form. Everyone would benefit if all cooperate, but only a very trusting fool would think it rational to cooperate himself. Perhaps some broader generalization to the many-person case would also be useful, but that matter will not be taken up here.

The cooperative outcome in the Stag Hunt can be assured by many of the same means as are discussed here for the PD. As might be expected, cooperation is somewhat easier to come by in the two person Stag Hunt than in the two person PD. Details will not be given here, but the interested reader may consult Skyrms 2004, which is responsible for a recent resurgence of interest in this game.

9. Asynchronous Moves

It has often been argued that rational self-interested players can obtain the cooperative outcome by making their moves conditional on the moves of the other player. Peter Danielson, for example, favors a strategy of *reciprocal cooperation*: if the other player would cooperate if you cooperate and would defect if you don't, then cooperate, but otherwise defect. Conditional strategies like this are ruled out in the versions of the game described above, but they may be possible in versions that more accurately model real world situations. In this section and the next, we consider two such versions. Here we eliminate the requirement that the two players move simultaneously. Consider the situation of a firm whose sole competitor has just lowered prices. Or suppose the buyer of a car has just paid the agreed purchase price and the seller has not yet handed over the title. We can think of these as situations in which one player has to choose to cooperate or defect after

the other player has already made a similar choice. The corresponding game is an *asynchronous* or *extended* PD.

Careful discussion of an asynchronous PD example, as Skyrms (1998) and Vanderschraaf recently note, occurs in the writings of David Hume, well before Flood and Dresher's formulation of the ordinary PD. Hume writes about two neighboring grain farmers:

> Your corn is ripe today; mine will be so tomorrow. 'Tis profitable for us both, that I shou'd labour with you to-day, and that you shou'd aid me to-morrow. I have no kindness for you, and know you have as little for me. I will not, therefore, take any pains on your account; and should I labour with you upon my own account, in expectation of a return, I know I shou'd be disappointed, and that I shou'd in vain depend upon your gratitude. Here then I leave you to labour alone: You treat me in the same manner. The seasons change; and both of us lose our harvests for want of mutual confidence and security.

In deference to Hume, Skyrms and Vanderschraaf refer to this kind of asynchronous PD as the "farmer's dilemma." It is instructive to picture it in a tree diagram.
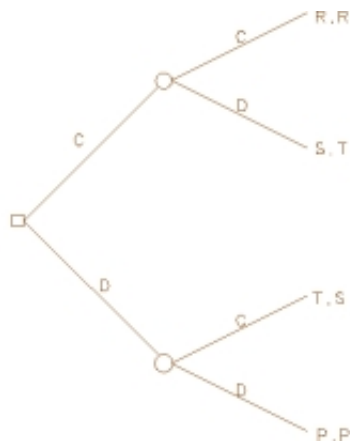

Figure 2

Here, time flows to the right. The node marked by a square indicates Player One's choice point, those marked by circles indicate Player Two's. The moves and the payoffs to each player are exactly as in the ordinary PD, but here Player Two can choose his move according to what Player One does. Tree diagrams like Figure 2 are said to be *extensive-form* game representations, whereas the payoff matrices given previously are *normal-form* representations. As Hume's analysis indicates, making the game asynchronous does not remove the dilemma. Player One knows that if he were to choose **C** on the first move, Player Two would choose **D** on the second move (since she prefers the temptation to the reward), so he would himself end up with the sucker payoff. If Player One were to choose **D**, Player Two would still choose **D** (since she prefers the punishment to the sucker payoff), and he would end up with the punishment payoff. Since he prefers the punishment

payoff to the sucker payoff, Player One will choose **D** on the first move and both players will end up with the punishment payoff. This kind of "backward" reasoning, in which the players first evaluate what would happen on the last move if various game histories were realized, and use this to determine what would happen on preceding moves applies quite broadly to games in extensive form, and a more general version of it will be discussed under *finite iteration* below.

The farmer's dilemma can be represented in normal form by understanding Player One to be choosing between **C** and **D** and Player Two to be (simultaneously) choosing among four conditional moves: cooperate unconditionally (**Cu**), defect unconditionally (**Du**), imitate Player One's move (**I**), and do the opposite of Player One's move (**O**). The result is a two player game with the following matrix.

	Cu	**Du**	**I**	**O**
C	R, R	S, T	R, R	S, T
D	T, S	P, P	P, P	T, S

The reader may note that this game is a (multiple-move) equilibrium dilemma. The sole (weak) nash equilibrium results when Player One chooses **D** and Player Two chooses **Du**, thereby achieving for themselves the inferior payoffs of P and P. The game is not, however, a dominance PD. Indeed, there is no dominant move for either player. It is commonly believed that rational self-interested players will reach a nash equilibrium even when neither player has a dominant move. If so, the farmer's dilemma is still a dilemma.

To preserve the symmetry between the players that characterizes the ordinary PD, we may wish to modify the asynchronous game. Let us take extended PD to be played in stages. First each player chooses a first move (**C** or **D**) and a second move (**Cu**, **Du**, **I** or **O**). Next a referee determines who moves first, giving each player an equal chance. Finally the outcome is computed in the appropriate way. For example, suppose Row plays (**D**, **O**) (meaning that he will defect if he moves first and do the opposite of his opponent if he moves second) and Column plays (**C**, **Du**). Then Row will get P if he goes first and T if he goes second, which implies that his expected payoff is $1/2(P+T)$. Column will get S if she goes first and P if she goes second, giving her an expected payoff of $1/2(P+S)$. It is straightforward, but tedious, to calculate the entire eight by eight payoff matrix. After doing so, the reader may observe that, like the farmer's dilemma, the symmetric form of the extended PD is an equilibrium PD, but not a dominance PD. The sole nash equilibrium occurs when both players adopt the strategy (**D**, **Du**), thereby achieving the inferior payoffs of (P, P).

It may be worth noting that an asynchronous version of the Stag Hunt, unlike the PD, presents few issues of interest. If the first player does his part in the hunt for stag on day one, the second should do her part on day two. If he hunts hare on day one, she should do likewise on day two. The first player, realizing this, should hunt stag on day one. So rational players should have no difficulty reaching the cooperative outcome in the asynchronous Stag Hunt.

10. Transparency

Another way that conditional moves can be introduced into the PD is by assuming that players have the property that David Gauthier has labeled *transparency*. A fully transparent player is one whose intentions are completely visible to others. Nobody holds that we humans are fully transparent, but the observation that we can often successfully predict what others will do suggests that we are at least "translucent." Furthermore agents of larger scale, like firms or countries, which may have to publicly deliberate before acting, may be more transparent than we are. Thus there may be some theoretical interest in investigations of PDs with transparent players. Such players could presumably execute conditional strategies more sophisticated than those of the (non- transparent) extended game players, strategies, for example that are conditional on the conditional strategies employed by others. There is some difficulty, however, in determining exactly what strategies are feasible for such players. Suppose Row adopted the strategy "do the same as Column" and Column adopted the strategy "do the opposite of Row". There is no way that both these strategies could be satisfied. On the other hand, if each adopted the strategy "imitate the other player", there are two ways the strategies could be satisfied, and there is no way to determine which of the two they would adopt. Nigel Howard, who was probably the first to study such conditional strategies systematically, avoided this difficulty by insisting on a rigidly typed hierarchy of games. At the base level we have the ordinary PD game, where each player chooses between **C** and **D**. For any game G in the hierarchy we can generate two new games RG and CG. In RG, Column has the same moves as in game G and Row can choose any function that assigns **C** or **D** to each of Column's possible moves. Similarly in CG, Row has the same moves as in G and Column has a new set of conditional moves. For example, if [PD] is the base level game, then C[PD] is the game in which Column can choose from among the strategies **Cu**, **Du**, **I** and **O** mentioned above. Howard observed that in the two third level games RC[PD] and CR[PD] (and in every higher level game) there is an equilibrium outcome giving each player R. In particular, such an equilibrium is reached when one player plays **I** and the other cooperates when his opponent plays **I** and defects when his opponent plays **Cu**, **Du** or **O**. Notice that this last strategy is tantamount to Danielson's *reciprocal cooperation*.

The lesson of all this for rational action is not clear. Suppose two players in a PD were sufficiently transparent to employ the conditional strategies of higher level games. How do

they decide what level game to play? Who chooses the imitation move and who chooses reciprocal cooperation? To make a move in a higher level game is presumably to form an intention observable by the other player. But why should either player expect the intention to be carried out if there is benefit in ignoring it?

Conditional strategies have a more convincing application when we take our inquiry as directed, not towards playing the PD, but as designing agents who would play it well with a variety of likely opponents. This is the viewpoint of Danielson. (See also J.V. Howard for an earlier enlightening discussion of this viewpoint.) A conditional strategy is not an intention that a player forms as a move in a game, but a deterministic algorithm defining a kind of player. Indeed, one of the lessons of the PD may be that transparent agents are better off if they can form irrevocable "action protocols" rather than always following the intentions they may form at the time of action. Danielson does not limit himself *a priori* to strategies within Howard's hierarchy. An agent is simply a computer program, which can contain lines permitting other programs to read and execute it. We could easily write two such programs, each designed to determine whether its opponent plays **C** or **D** and to do the opposite. What happens when these two play a PD depends on the details of implementation, but it is likely that they will be "incoherent," i.e., they will enter endless loops and be unable to make any move at all. To be successful a program *should* be able to move when paired with a variety of other programs, including copies of itself, and it should be able to get valuable outcomes. Programs implementing **I** and **O** in a straightforward way are not likely to succeed because when paired with each other they will be incoherent. Programs implementing **Du** are not likely to succeed because they get only P when paired with their clones. Those implementing **Cu** are not likely to succeed because they get only S when paired with programs that recognize and exploit their unconditionally cooperative nature. There is some vagueness in the criteria of success. In Howard's scheme we could compare a conditional strategy with all the possible alternatives of that level. Here, where any two programs can be paired, that approach is senseless. Nevertheless, certain programs seem to do well when paired with a wide variety of players. One is a version of the strategy that Gauthier has advocated as *constrained maximization*. The idea is that a player *j* should cooperate if the other would cooperate if *j* did, and defect otherwise. As stated, this appears to be a strategy for the *RC*[PD] or *CR*[PD] games. It is not clear how a program implementing it would move (if indeed it does move) when paired with itself. Danielson is able to construct an approximation to *constrained maximization*, however, that does cooperate with itself. Danielson's program (and other implementations of *constrained maximization*) cannot be coherently paired with everything. Nevertheless it does move and score well against familiar strategies. It cooperates with **Cu** and itself and it defects against **Du**. If it is coherently paired it seems guaranteed a payoff no worse than P.

A second successful program models Danielson's *reciprocal cooperation*. Again, it is not clear that the strategy (as formulated above) allows it to cooperate (or make any move)

with itself, but Danielson is able to construct an approximation that does. The (approximate) *reciprocal cooperation* does as well as (approximate) *constrained maximization* against itself, **Du** and *constrained maximization*. Against **Cu** it does even better, getting *T* where *constrained maximization* got only *R*.

11. Finite Iteration

Many of the situations that are alleged to have the structure of the PD, like defense appropriations of military rivals or price setting for duopolistic firms are better modeled by an iterated version of the game in which players play the PD repeatedly, retaining access at each round to the results of all previous rounds. In these iterated PDs (hence forth IPDs) players who defect in one round can be "punished" by defections in subsequent rounds and those who cooperate can be rewarded by cooperation. Thus the appropriate strategy for rationally self-interested players is no longer obvious. The theoretical answer to this question, it turns out, depends strongly on the definition of IPD employed and the knowledge attributed to rational players.

An IPD can be represented in extensive form by a tree diagram like the one for the farmer's dilemma above.
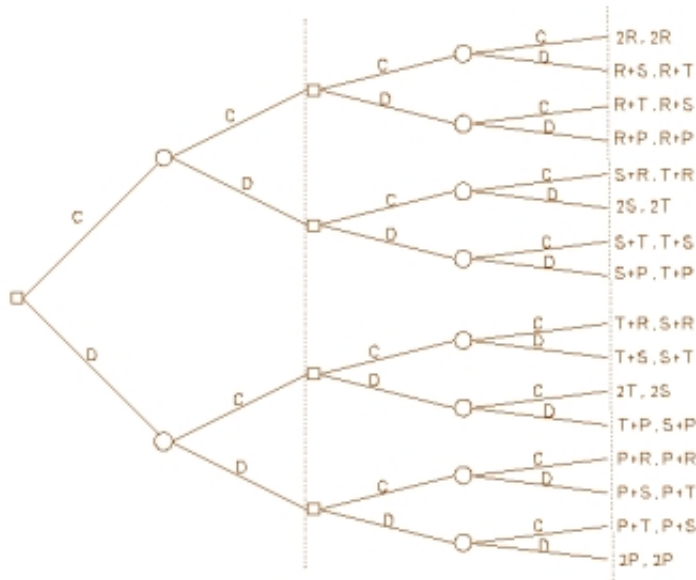

Figure 3

Here we have an IPD of length two. The end of each of the two rounds of the game is marked by a dotted vertical line. The payoffs to each of the two players (obtained by adding their payoffs for the two rounds) are listed at the end of each path through the tree. The representation differs from the previous one in that the two nodes on each branch within the same division mark simultaneous choices by the two players. Since neither player knows the move of the other at the same round, the IPD does not qualify as one of

the game theorist's standard "games of perfect information." If the players move in succession rather than simultaneously (which we might indicate by removing the dotted vertical lines), the resulting game is an iterated farmer's dilemma, which does meet the game theorist's definition and which shares many of the features that make the IPD interesting.

Like the farmer's dilemma, an IPD can, in theory, be represented in normal form by taking the players' moves to be *strategies* telling them how to move if they should reach any node at the end of a round of the game tree. The number of strategies increases very rapidly with the length of the game so that it is impossible in practice to write out the normal form for all but the shortest IPD's. Every pair of strategies determines a "play" of the game, i.e., a path through the extensive-form tree.

In a game like this, the notion of nash equilibrium loses some of its privileged status. Recall that a pair of moves is a nash equilibrium if each is a best reply to the other. Let us extend the notation used in the discussion of the asynchronous PD and let **Du** be the strategy that calls for defection at every node of an IPD. It is easy to see that **Du** and **Du** form a nash equilibrium. But against **Du**, a strategy that calls for defection unless the other player cooperated at, say, the fifteenth node, would determine the same play (and therefore the same payoffs) as **Du** itself does. The components that call for cooperation never come into play, because the other player does not cooperate on the fifteenth (or any other) move. Similarly, a strategy calling for cooperation after the second cooperation by itself does equally well. Thus these strategies and many others form nash equilibria with **Du**. There is a sense in which these strategies are clearly not equally rational. Although they yield the same payoffs at the nodes along the path representing the actual play, they would not yield the same payoffs if other nodes had been reached. If Player One *had* cooperated in the past, that would still provide no good reason for him to cooperate now. A nash equilibrium requires only that the two strategies are best replies to each other as the game actually develops. A stronger solution concept for extensive-form games requires that the two strategies would still be best replies to each other no matter what node on the game tree were reached. This notion of *subgame-perfect equilibrium* is defined and defended in Selten 1975. It can be expressed by saying that the strategy-pair is a nash equilibrium for every subgame of the original game, where a subgame is the result of taking a node of the original game tree as the root, pruning away everything that does not descend from it.

Given this new, stronger solution concept, we can ask about the solutions to the IPD. There is a significant theoretical difference on this matter between IPDs of fixed, finite length, like the one pictured above, and those of infinite or indefinitely finite length. In games of the first kind, one can prove by an argument known as *backward induction* that **Du**, **Du** is the only subgame perfect equilibrium. Suppose the players know the game will last exactly n rounds. Then, no matter what node have been reached, at round n-1 the players face an

ordinary ("one-shot") PD, and they will defect. At round *n*-2 the players know that, whatever they do now, they will both defect at the next round. Thus it is rational for them to defect now as well. By repeating this argument sufficiently many times, the rational players deduce that they should defect at every node on the tree. Indeed, since at every node defection is a best response to any move, there can be no other subgame-perfect equilibria.

In practice, there is not a great difference between how people behave in long fixed-length IPDs (except in the final few rounds) and those of indeterminate length. This suggests that some of the rationality and common knowledge assumptions used in the backward induction argument (and elsewhere in game theory) are unrealistic. There is a considerable literature attempting to formulate the argument carefully, examine its assumptions, and to see how relaxing unrealistic assumptions might change the rationally acceptable strategies in the PD and other games of fixed length. (For a small sample, see Bovens, Kreps *et al*, Kreps and Wilson, Pettit and Sugden, Sobel 1993 and Binmore 1997).

Player One's belief that there is a slight chance that Two might pursue an "irrational" strategy other than continual defection could make it rational for her to cooperate frequently herself. Indeed, even if One were certain of Two's rationality, One's belief that there was some chance that Two believed she harbored such doubts could have the same effect. Thus the argument for continual defection in the IPD of fixed length depends on complex iterated claims of certain knowledge of rationality. An even more unrealistic assumption, noted by Rabinowicz and others, is that each player continue to believe that the other will choose rationally on the next move even after evidence of irrational play on previous moves. For example, it is assumed that, at the node reached after a long series of moves (**C**, **C**), …,(**C**, **C**), Player One will choose **D** despite never having done so before.

Some have used these kinds of observation to argue that the backward induction argument shows that standard assumptions about rationality (with other plausible assumptions) are inconsistent or self-defeating. For (with plausible assumptions) one way to ensure that a rational player will doubt one's own rationality is to behave irrationally. In the fixed-length IPD, for example, Player One may be able to deduce that, if she were to follow an appropriate "irrational" strategy, Player Two would rationally react so that they can achieve mutual cooperation in almost all rounds. So our assumptions seem to imply both that Player One should continually defect and that she would do better if she didn't. (See Skyrms 1990, pp. 125-139 and Bicchieri 1989.)

12. The Centipede and the Finite IPD

Many of the issues raised by the fixed-length IPD can be raised in even starker form by a somewhat simpler game. Consider a PD in which they punishment payoff is zero. Now

iterate the asynchronous version of this game a fixed number times. Imagine that both players are restricted to highly "punitive" strategies according to which, they must always defect against a player who has ever defected. (One important strategy of this variety is discussed below under the label **GRIM**.) The result is a *centipede* game. A particularly nice realization is given by Sobel 2005. A stack of *n* one-dollar bills lies on a table. Players take turns taking money from the stack, one or two bills per turn. The game ends when the stack runs out or one of the players takes two bills (whichever comes first). Both players keep what they have taken to that point. The extensive form of the game with for *n*=4 is pictured below.
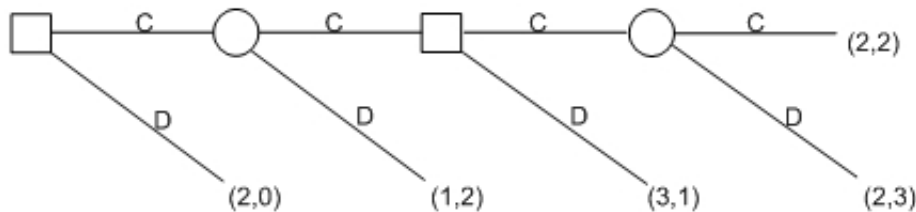

Figure 4

Presumably the true centipede would contain 100 "legs" and the general form discussed here should really be called the "*n*-tipede." The game appears to be discussed first in Rosenthal.

As in the fixed-length PD, a backward induction argument easily establishes that a rational player should take two bills on his first move, giving her a payoff of $2 or $3 dollars, depending on whether she moves first or second, and leaving the remainder of the *n* dollars undistributed. In more technical terms, the only nash equilibria of the game are those where the first player takes two dollars on the first move and the only subgame perfect equilibrium is the one in which both players take $2 on any turn they should get. Again, common sense and experimental evidence suggest that real players rarely act in this way and this leads to questions about exactly what assumptions this kind of argument requires and whether they are realistic. (In addition to the sample mentioned in the section on finitely iterated PDs, see, for example, Aumann 1998, Selten 1978, and Rabinowicz.) The centipede also raises the same questions about cooperation and socially desirable altruism as does the PD and it is a favorite tool in empirical investigations of game playing.

13. Infinite Iteration

One way to avoid the dubious conclusion of the backward induction argument without

delving too deeply into conditions of knowledge and rationality is to consider infinitely repeated PDs. No human agents can actually play an infinitely repeated game, of course, but the infinite IPD has been considered an appropriate way to model a series of interactions in which the participants never have reason to think the current interaction is their last. In this setting a pair of strategies determines an infinite path through of the game tree. If the payoffs of the one-shot game are positive, their total along any such path is infinite. This makes it somewhat awkward to compare strategies. If we confine ourselves to those strategies that can be implemented by mechanical devices (with finite memories and speeds of computation), however, it turns out that the sequence of payoffs to each player will always, after a finite number of rounds, cycle repeatedly through a particular finite sequence of payoffs. The relative value of such infinite sequences of payoffs can then be identified with the average value of the payoffs in one cycle. This value reflects the limit of average payoff per round as the number of rounds increases. (See Binmore 1992, page 365 for further justification.) Since there is no last round, it is obvious that backward induction does not apply to the infinite IPD.

14. Indefinite Iteration

Most contemporary investigations the IPD take it to be neither infinite nor of fixed finite length but rather of indeterminate length. This is accomplished by including in the game specification a probability p (the "shadow of the future") such that at each round the game will continue with probability p. Alternatively, a "discount factor" p is applied to the payoffs after each round so that present payoffs are valued more highly than future. Mathematically, it makes little difference whether p is regarded as a probability of continuation or a discount on payoffs. The value of cooperation at a given stage in an IPD clearly depends on the odds of meeting one's opponent in later rounds. (This has been said to explain why the level of courtesy is higher in a village than a metropolis and why customers tend leave better tips in local restaurants than distant ones.) As p approaches zero, the IPD becomes a one-shot PD, and the value of defection increases. As p approaches one the IPD becomes an infinite IPD, and the value of defection decreases. It is also customary to insist that the game has the property labeled RCA above, so that (in the symmetric game) players do better by cooperating on every round than they would do by "taking turns" — you cooperate while I defect and then I cooperate while you defect.

There is an observation, apparently originating in Kavka, 1983, and given more mathematical form in Carroll, that the backward induction argument applies as long as an upper bound to the length of the game is common knowledge. For if b is such an upper bound, then, if the players were to get to stage b, they would know that it was the last round and they would defect; if they were to get to stage b-1, they would know that their behavior on this round cannot affect the decision to defect on the next, and so they would defect; and so on. It seems an easy matter to compute upper bounds on the number of

interactions in real-life situations. For example, since shopkeeper Jones cannot make more than one sale a second and since he will live less than a thousand years, he and customer Smith can calculate (conservatively) that they cannot possibly conduct more than 10^{12} transactions. It is instructive to examine this argument more closely in order to dramatize the assumptions made in standard treatments of the indefinite IPD and other indefinitely repeated games. Note first that, in an indefinite IPD as described above, there can be no upper bound on the length of the game. There is, instead, some fixed probability p that, at any time in which the game is still being played, it will continue to be played with probability p. If the interaction of Smith and Jones were modeled as an indefinite IPD, therefore, the probability of their interacting in a thousand years would not be zero, but rather some number greater than p^k where p is the probability of their interacting again now and k is the number of seconds in a thousand years. A more realistic way to model the interaction might be to allow the value of p to decrease as the game progressed. As long as p always remains greater than zero, however, it remains true that there can be no upper bound on the number of possible interactions, i.e., no time at which the probability of future interactions becomes zero. Suppose, on the other hand, that there *was* a number n such that that there was zero probability of the game's continuing to stage n. Let $p_1, \ldots, p_n$ be the probabilities that game continues after stage 1, …, stage n. Then there must be a smallest i such that p_i becomes 0. (It would happen at $i=n$ if not sooner.) Given the standard common knowledge assumptions that we have been making, the players would know this value of i, and the IPD would be one of fixed length, and not an indefinite IPD at all. In the case of the shopkeeper and his customer, we are to suppose that both know today that their last interaction will occur, let's say, at noon on June 10th, 2020. The very plausible idea that we began with, viz., that *some* upper bounds on the number of interactions are common knowledge, even though the smallest upper bound is not, is incompatible with the assumption that we know all the continuation probabilities p_i from the start.

As Becker and Cudd astutely observe, we don't need an upper bound on the number of possible iterations to make a backward induction argument for defection possible. If the players know all the values of p_i from the outset, then, as long as the value of p_i becomes and remains sufficiently small, they (and we) can compute a stage k at which the risk of future punishment and the chance of future reward no longer outweighs the benefit of immediate defection. So they know their opponent will defect at stage k, and the induction begins. This modification of the Kavka/Carroll argument, however, only further exposes the implausibility of its assumptions. Not only are Smith and Jones expected to believe that there is non-zero probability that they will be interacting in a thousand years, each is expected to be able to compute the precise day on which future interactions will become and remain so unlikely that their expected future return is outweighed by that day's payoff. Furthermore each is expected to believe that the other has made this computation, and that the other expects him to have made it, and so on.

The iterated version of the PD was discussed from the time the game was devised, but interest accelerated after influential publications of Robert Axelrod in the early eighties. Axelrod invited professional game theorists to submit computer programs for playing IPDs. All the programs were entered into a tournament in which each played every other (as well as a clone of itself and a strategy that cooperated and defected at random) hundreds of times. It is easy to see that in a game like this no strategy is "best" in the sense that its score would be highest among any group of competitors. If the other strategies never consider the previous history of interaction in choosing their next move, it would be best to defect unconditionally. If the other strategies all begin by cooperating and then "punish" any defection against themselves by defecting on all subsequent rounds, then a policy of unconditional cooperation is better. Nevertheless, as in the transparent game, some strategies have features that seem to allow them to do well in a variety of environments. The strategy that scored highest in Axelrod's initial tournament, Tit for Tat (henceforth **TFT**), simply cooperates on the first round and imitates its opponent's previous move thereafter. More significant than **TFT**'s initial victory, perhaps is the fact that it won Axelrod's second tournament, whose sixty three entrants were all given the results of the first tournament. In analyzing the his second tournament, Axelrod noted that each of the entrants could be assigned one of five "representative" strategies in such a way that a strategy's success against a set of others can be accurately predicted by its success against their representative. As a further demonstration of the strength of **TFT**, he calculated the scores each strategy would have received in tournaments in which one of the representative strategies was five times as common as in the original tournament. **TFT** received the highest score in all but one of these hypothetical tournaments.

Axelrod attributed the success of **TFT** to four properties. It is *nice*, meaning that it is never the first to defect. The eight nice entries in Axelrod's tournament were the eight highest ranking strategies. It is *retaliatory*, making it difficult for it to be exploited by the rules that were not nice. It is *forgiving*, in the sense of being willing to cooperate even with those who have defected against it (provided their defection wasn't in the immediately preceding round). An unforgiving rule is incapable of ever getting the reward payoff after its opponent has defected once. And it is *clear*, presumably making it easier for other strategies to predict its behavior so as to facilitate mutually beneficial interaction.

Suggestive as Axelrod's discussion is, it is worth noting that the ideas are not formulated precisely enough to permit a rigorous demonstration of the supremacy of **TFT**. One doesn't know, for example, the extent of the class of strategies that might have the four properties outlined, or what success criteria might be implied by having them. It is true that if one's opponent is playing **TFT** (and the shadow of the future is sufficiently large) then one's maximum payoff is obtained by a strategy that results in mutual cooperation on every round. Since **TFT** is itself one such strategy this implies that **TFT** forms a nash equilibrium with itself in the space of all strategies. But that does not particularly

distinguish **TFT**, for **Du**, **Du** is also a nash equilibrium. Indeed, a "folk theorem" of iterated game theory (now widely published — see, for example, Binmore 1992, pp. 373-377) implies that, for any p, $0 \le p \le 1$ there is a nash equilibrium in which p is the fraction of times that mutual cooperation occurs. Indeed **TFT** is, in some respects, *worse* than many of these other equilibrium strategies, because the folk theorem can be sharpened to a similar result about subgame perfect equilibria. **TFT** is, in general, *not* subgame perfect. For, were one **TFT** player (*per impossible*) to defect against another in a single round, the second would have done better as an unconditional cooperator.

15. Iteration With Error

In a survey of the field several years after the publication of the results reported above, Axelrod and Dion, chronicle several successes of **TFT** and modifications of it. They conclude that "research has shown that many of Axelrod's findings…can be generalized to settings that are quite different from the original two-player iterated Prisoner's Dilemma game." But in several reasonable settings **TFT** has serious drawbacks. One such case, noted in the Axelrod and Dion survey, is when attempts are made to incorporate the plausible assumption that players are subject to errors of execution and perception. There are a number of ways this can be done. Bendor, for example, considers "noisy payoffs." When a player cooperates while its opponent defects, its payoff is $S+e$, where e is a random variable whose expected value is 0. Each player infers the other's move from its own payoff, and so if e is sufficiently high its inference may be mistaken. Sugden (pp 112-115) considers players who have a certain probability of making an error of execution that is apparent to them but not their opponents. Such players can adopt strategies by which they "atone" for mistaken defections by being more cooperative on later rounds than they would be after intended defection. Assuming that players themselves cannot distinguish a mistaken move or observation from a real one, however, the simplest way to model the inevitability of error is simply to forbid completely deterministic strategies like **TFT**, replacing them with "imperfect" counterparts, like "imitate the other player's last move with 99% probability and oppose it with 1% probability." **Imperfect TFT** is much less attractive than its deterministic sibling, because when two **imperfect TFT** strategies play each other, an "error" by either one will set off a long chain of moves in which the players take turns defecting. In a long iterated game between two **imperfect TFT**'s with any probability p of error, $0 < p < 1/2$, players will approach the same average payoffs as in a game between two strategies that choose randomly between cooperation and defection, namely $1/4(R+P+S+T)$. That is considerably worse than the payoff of R, that results when $p=0$.

The predominant view seems to be that, when imperfection is inevitable, successful strategies will have to be more forgiving of defections by their opponents (since those defections might well be unintended). Molander 1985 demonstrates that strategies that mix

TFT with **Cu** do approach a payoff of R as the probability of error approaches zero. When these mixes play each other, they benefit from higher ratios of **Cu** to **TFT**, but if they become too generous, they risk exploitation by "stingy" strategies that mix **TFT** with defection. Molander calculates that when the mix is set so that, following a defection, one cooperates with probability $g(R, P, T, S)=\min\{1-(T-R)/(R-S), (R-P)/(T-P)\}$, the generous strategies will get the highest score with each other that is possible without allowing stingy strategies to do better against them than TFT does. Following Nowak and Sigmund, we label this strategy **generous TFT**, or **GTFT**.

The idea that the presence of imperfection induces greater forgiveness or generosity is only plausible for low levels of imperfection. As the level of imperfection approaches 1/2, **Imperfect TFT** becomes indistinguishable from the random strategy, for which the very ungenerous **Du** is the best reply. A simulation by Kollock seems to confirm that at high levels of imperfection, more stinginess is better policy than more forgiveness. But Bendor, Kramer and Swistak note that the strategies employed in the Kollock simulation are not representative and so the results must be interpreted with caution.

A second idea is that an imperfect environment encourages strategies to observe their opponent's play more carefully. In a tournament similar to Axelrod's (Donninger) in which each player's moves were subject to a 10% chance of alteration, **TFT** finished sixth out twenty-one strategies. As might have predicted on the dominant view, it was beaten by the more generous **Tit-for-Two-Tats** (which cooperates unless defected against twice in a row). It was also beaten, however, by two versions of **Downing**, a program that bases each new move on its best estimate how responsive its opponent has been to its previous moves. In Axelrod's two original tournaments, **Downing** had ranked near the bottom third of the programs submitted. Bendor (1987) demonstrates deductively that against imperfect strategies there are advantages to basing one's probability of defection on longer histories than does **TFT**.

One clever implementation of the idea that a strategies in an imperfect environment should pay attention to their previous interactions is the family of "Pavlovian" strategies investigated by Kraines and Kraines. For each natural number n, n-Pavlov, or $\mathbf{P}_n$, adjusts its probability of cooperation in units of $1/n$, according to how well it fared on the previous round. More precisely, if $\mathbf{P}_n$ was cooperating with probability p on the last round, then on this round it will cooperate with probability $p[+](1/n)$ if it received the reward payoff on the previous round, $p[-](1/n)$ if it received the punishment payoff, $p[+](2/n)$if it received the temptation payoff, and $p[-](2/n)$ if it received the sucker payoff. [+] and [−] are bounded addition and subtraction, i.e., $x[+]y$ is the sum $x+y$ unless that number exceeds one, in which case it is one (or as close to one as the possibility of error allows), and $x[-]y$ is similarly either $x-y$ or close to zero. Strictly speaking, $\mathbf{P}_n$ is not fully specified until an initial probability of cooperation is given, but for most purposes the value of that parameter

becomes insignificant in sufficiently long games and can be safely ignored. It may appear that $\mathbf{P}_n$ requires far more computational resources to implement than, say, **TFT**. Each move for the latter depends on only on its opponent's last move, whereas each move for $\mathbf{P}_n$ is a function of the entire history of previous moves of both players. $\mathbf{P}_n$, however, can always calculate its next move by tracking only its current probability of cooperation and its last payoff. As its authors maintain, this seems like "a natural strategy in the animal world." One can calculate that for $n>1$, $\mathbf{P}_n$ does better against the random strategy than does **TFT**. More generally, $\mathbf{P}_n$ does as well or better than **TFT** against the generous unresponsive strategies **Cp** that always cooperate with fixed probability $p \geq 1/2$ (because an occasional temptation payoff can teach it to exploit the unresponsive strategies.) In these cases the "slow learner" versions of Pavlov with higher values of n do slightly better than the "fast learners" with low values. Against responsive strategies, like other Pavlovian strategies and **TFT**, $\mathbf{P}_n$ and its opponent eventually reach a state of (almost) constant cooperation. The total payoff is then inversely related to the "training time," i.e., the number of rounds required to reach that state. Since training time of $\mathbf{P}_n$ varies exponentially with n, Kraines and Kraines maintain that $\mathbf{P}_3$ or $\mathbf{P}_4$ are to be preferred to other Pavlovian strategies, and are close to "ideal" IPD strategies. It should be noted, however, that when (deterministic) **TFT** plays itself, no training time at all is required, whereas when a Pavlovian strategy plays **TFT** or another Pavlov, the training time can be large. Thus the cogency of the argument for the superiority of Pavlov over **TFT** depends on the observation that its performance shows less degradation when subject to imperfections. It is also worth remembering that no strategy is best in every environment, and the criteria used in defense of various strategies in the IPD are vague and heterogeneous. One advantage of the evolutionary versions of the IPD discussed in the next section is that they permit more careful formulation and evaluation of success criteria.

16. Evolution

Perhaps the most active area of research on the PD concerns evolutionary versions of the game. A population of players employing various strategies play IPDs among themselves. The lower scoring strategies decrease in number, the higher scoring increase, and the process is repeated. Thus success in an evolutionary PD (henceforth EPD), requires doing well with other successful strategies, rather than doing well with a wide range of strategies.

The initial population in an EPD can be represented by a set of pairs $\{(p_1, s_1), \ldots (p_n, s_n)\}$ where $p_1, \ldots, p_n$ are the proportions of the population playing strategies $\mathbf{s_1}, \ldots, \mathbf{s_n}$, respectively. The description of EPDs given above does not specify exactly how the population of strategies is to be reconstituted after each IPD. The usual assumption, and the most sensible one for biological applications, is that a score in any round indicates the relative number of "offspring" in the next. It is assumed that the size of the entire population stays fixed, so that births of more successful strategies are exactly offset by

deaths of less successful ones. This amounts to the condition that the proportion $p_i{*}$ of each strategy s_i in the successor population is determined by the equation $p_i{*}=p_i(V_i/V)$, where V_i is the score of s_i in the previous round and V is the average of all scores in the population. Thus every strategy that scores above the population average will increase in number and every one that scores below the average will decrease. This kind of evolution is referred to as "replicator dynamics" or evolution according to the "proportional fitness" rule. Other rules of evolution are possible. Bendor and Swistak argue that, for social applications, it makes more sense to think of the players as switching from one strategy to another rather than as coming into and of existence. Since rational players would presumably switch only to strategies that received the highest payoff in previous rounds, only the highest scoring strategies would increase in numbers. A variety of other possible evolutionary dynamics are described and discussed in Kuhn 2004. Discussion here, however, will primarily concern EPDs with the proportional fitness rule.

Axelrod, borrowing from Trivers and Maynard Smith, includes a description of the EPD with proportional fitness, and a brief analysis of the evolutionary version of his IPD tournament. For Axelrod, the EPD provides one more piece of evidence in favor of **TFT**:

> TIT FOR TAT had a very slight lead in the original tournament, and never lost this lead in simulated generations. By the one-thousandth generation it was the most successful rule and still growing at a faster rate than any other rule.

Axelrod's EPD tournament, however, incorporated several features that might be deemed artificial. First, it permitted deterministic strategies in a noise-free environment. As noted above, **TFT** can be expected to do worse under conditions that model the inevitability of error. Second, it began with only the 63 strategies from the original IPD tournament. Success against strategies concocted in the ivory tower may not imply success against all those that might be found in nature. Third, the only strategies permitted to compete at a given stage were the survivors from the previous stage. A more realistic model, one might argue, would allow new "mutant" strategies to enter the game at any stage. Changing this third feature might well be expected to hurt **TFT**. For a large growth in the **TFT** population would make it possible for mutants employing more naive strategies like **Cu** to regain a foothold, and the presence of these naifs in the population might favor nastier strategies like **Du** over **TFT**.

Nowak and Sigmund simulated two kinds of tournaments that avoid the three questionable features. The first examined the family of "reactive" strategies. For any probabilities $y, p,$ and q, $\mathbf{R}(y, p, q)$ is the strategy of cooperating with probability y in the first round and thereafter with probability p if the other player has cooperated in the previous round, and with probability q if she has defected. This is a broad family, including many of the strategies already considered. **Cu**, **Du**, **TFT**, and **Cp** are $\mathbf{R}(1,1,1)$, $\mathbf{R}(0,0,0)$, $\mathbf{R}(1,1,0)$, and

$\mathbf{R}(p, p, p)$. To capture the inevitability of error, Nowak and Sigmund exclude the deterministic strategies, where p and q are exactly 1 or 0, from their tournaments. As before, if the game is sufficiently long (and p and q are not integers), the first move can be ignored and a reactive strategy can be identified with its p and q values. Particular attention is paid to the strategies close to Molander's **GTFT** described above, where $p=1$ and $q=\min\{1-(T\text{-}R)/(R\text{-}S),\ (R\text{-}P)/(T\text{-}P)\}$. The first series of Nowak and Sigmund's EPD tournaments begin with representative samples of reactive strategies. For most such tournaments, they found that evolution led irreversibly to **Du**. Those strategies $\mathbf{R}(p, q)$ closest to $\mathbf{R}(0,0)$ thrived while the others perished. When one of the initial strategies is very close to **TFT**, however, the outcome changes.

> TFT and all other reciprocating strategies (near $(1,0)$) seem to have disappeared. But an embattled minority remains and fights back. The tide turns when 'suckers' are so decimated that exploiters can no longer feed on them. Slowly at first, but gathering momentum, the reciprocators come back, and the exploiters now wane. But the **TFT**-like strategy that caused this reversal of fortune is not going to profit from it: having eliminated the exploiters, it is robbed of its mission and superseded by the strategy closest to **GTFT**. Evolution then stops. Even if we introduce occasionally 1% of another strategy it will vanish.

On the basis of their tournaments among reactive strategies, Nowak and Sigmund conjectured that, while **TFT** is essential for the emergence of cooperation, the strategy that actually underlies persistent patterns of cooperation in the biological world is more likely to be **GTFT**.

A second series of simulations with a wider class of strategies, however, forced them to revise their opinion. The strategies considered in the second series allowed each player to base its probability of cooperation on its own previous move as well as its opponent's. A strategy can now be represented as $\mathbf{S}(p_1, p_2, p_3, p_4)$ where p_1, p_2, p_3, p_4 are the probabilities of defecting after outcomes $(\mathbf{C}, \mathbf{C})$, $(\mathbf{C}, \mathbf{D})$, $(\mathbf{D}, \mathbf{C})$, and $(\mathbf{D}, \mathbf{D})$, respectively i.e., after receiving the reward, sucker, temptation and punishment payoffs. (Again, we can ignore the probability of defecting on the first move as long as the p_is are not zero or one.) The initial population in these tournaments all play the random strategy $\mathbf{S}(.5, .5, .5, .5)$ and after every 100 generations a small amount of a randomly chosen (non-deterministic) mutant is introduced, and the population evolves by proportional fitness. The results are quite different than before. After 10^7 generations, a state of steady mutual cooperation was reached in 90% of the simulation trials. But less than 8.3% of these states were populated by players using **TFT** or **GTFT**. The remaining 91.7% were dominated by strategies close to $\mathbf{S}(1,0,0,1)$. This is the just the Pavlovian strategy $\mathbf{P}_1$ of Kraines and Kraines, which cooperates after receiving R or T and defects after receiving P or S. Kraines and Kraines

had been somewhat dismissive of P_1. They recall that Rapoport and Chammah, who identified it early in the history of game theory had labeled it "simpleton" and remark that "the appellation is well deserved". Indeed, P_1 has the unfortunate characteristic of trying to cooperate with **Du** on every other turn, and against **TFT** it can get locked into the inferior repeating series of payoffs $T, P, S, T, P, S, \ldots$. But Nowak and Sigmund's simulations suggest that these defects do not matter very much in evolutionary contexts. One reason may be that P_1 helps to make its environment unsuitable for its enemies. **Du** does well in an environment with generous strategies, like **Cu** or **GTFT**. **TFT**, as we have seen, allows these strategies to flourish, which could pave the way for **Du**. Thus, although **TFT** fares less badly against **Du** than P_1 does, P_1 is better at keeping its environment free of **Du**.

Simulations in a universe of deterministic strategies yield results quite different than those of Nowak and Sigmund. Bruce Linster (1992 and 1994) suggests that natural classes of strategies and realistic mechanisms of evolution can be defined by representing strategies as simple *Moore machines*. For example, P_1 is represented by the machine pictured below.

Figure 5

This machine has two states, indicated by circles. It begins in the leftmost state. The **C** in the left circle means that the machine cooperates on the first move. The arrow leading from the left to the right circle indicates that machine defects (enters the **D**) after it has cooperated (been in the **C** state) and its opponent has defected (the arrow is labeled by d). Linster has conducted simulations of evolutionary PD's among the strategies that can be represented by two-state Moore machines. It turns out that these are exactly the deterministic versions of the **S** strategies of Nowak and Sigmund. Since the strategies are deterministic, we must distinguish between the versions that cooperate on the first round and those that defect on the first round. Among the first round cooperators, **S**(0,0,0,0), **S**(0,0,0,1), **S**(0,0,1,0) and **S**(0,0,1,1) all represent the strategy **Cu** of unconditional cooperation. Similarly, four of the first-round defectors all represent **Du**. Each of the other **S**(p_1, p_2, p_3, p_4) where p_1, p_2, p_3, p_4 are either zero or one represent unique strategies. By deleting the six duplicates from the thirty-two deterministic versions of Nowak and

Sigmund's strategies, we obtain the twenty-six "two-state" strategies considered by Linster.

Linster simulated a variety of EPD tournaments among the two-state strategies. Some used "uniform mutation" in which each strategy in the population has an equal probability m of mutating into any of the other strategies. Some used "stylized mutation" in which the only mutations permitted are those that can be understood as the result of a single "broken link" in the Moore machine diagrams. In some, mutations were assumed to occur to a tiny proportion of the population at each generation; in others the "mutants" represented an invading force amounting to one percent of the original population. In some, a penalty was levied for increased complexity in the form of reduced payoffs for machines requiring more states or more links. As one might expect, results vary somewhat depending on conditions. There are some striking differences, however, between all of Linster's results and those of Nowak and Sigmund. In Linster's tournaments, no single strategy ever dominated the surviving populations in the way that P_1 and **GTFT** did in Nowak and Sigmund's. The one strategy that did generally come to comprise over fifty percent of the population was the initially-cooperating version of **S**(0,1,1,1). This is a strategy whose imperfect variants seem to have been remarkably uncompetitive for Nowak and Sigmund. It has been frequently discussed in the game theory literature under the label **GRIM** or **TRIGGER**. It cooperates until its opponent has defected once, and then defects for the rest of the game. According to Skyrms (1998) and Vanderschraaf, both Hobbes and Hume identified it as the strategy that underlies our cooperative behavior in important PD-like situations. The explanation for the discrepancy between **GRIM's** performance for Linster and for Nowak and Sigmund probably has to do with its sharp deterioration in the presence of error. In a match between two imperfect **GRIM**s, an "erronious" defection by either leads to a long string of mutual defections. Thus, in the long run imperfect **GRIM** does poorly against itself. The other strategies that survived (in lesser numbers) Linster's tournaments are **TFT**, P_1, **Cu**, and the initially-cooperative **S**(0,1,1,0). (Note that imperfect **GRIM** is also likely to do poorly against imperfect versions of these.) The observation that evolution might lead to a stable mix of strategies (perhaps each serving to protect others against particular types of invaders) rather than a single dominant strategy is suggestive. Equally suggestive is the result obtained under a few special conditions in which evolution leads to a recurring cycle of population mixes.

One might expect it to be possible to predict the strategies that will prevail in EPDs meeting various conditions, and to justify such predictions by formal proofs. Until recently, however, mathematical analyses of the EPD have been plagued by conceptual confusions about "evolutionary stability," the condition under which, as Nowak and Sigmund say, "evolution stops". Axelrod and Axelrod & Hamilton claim to show that **TFT** is evolutionarily stable. Selten 1983, includes an example of a game with no evolutionarily stable strategy, and Selten's argument that there is no such strategy clearly applies to the **EPD** and other evolutionary games. Boyd and Lorberbaum and Farrell and Ware present

still different proofs demonstrating that no strategies for the **EPD** are evolutionarily stable. Unsurprisingly, the paradox is resolved by observing that the three groups of authors each employ slightly different conceptions of evolutionary stability. The conceptual tangle is unraveled in a series of papers by Bendor and Swistak. Two stability concepts are described and applied to the EPD below. Readers who wish to compare these with some others that appear in the literature may consult the following brief guide:

Concepts of Stability in Evolutionary Games.

A strategy **s** for an evolutionary game has *universal strong narrow stability* ("usn-stability") if a population playing strategy **s** will, under any rule of evolution, drive to extinction any sufficiently small group of invaders all of which play the same strategy. An evolutionary game has usn-stability just in case it meets a simple condition on payoffs identified by Maynard Smith:

MS) For all strategies $\mathbf{j}$, $V(\mathbf{i}, \mathbf{i}) > V(\mathbf{j}, \mathbf{i})$ or both $V(\mathbf{i}, \mathbf{i})=V(\mathbf{j}, \mathbf{i})$ and $V(\mathbf{i}, \mathbf{j})>V(\mathbf{j}, \mathbf{j})$

(Here, and in what follows, the notation $V(\mathbf{i}, \mathbf{j})$ indicates the payoff to strategy **i** when **i** plays **j**.) MS says that any invaders do strictly worse against the natives than the natives themselves do against the natives or else they get exactly the same payoff against the natives as the natives themselves do, but the native does better against the invader than the invader himself does.

For any strategy **i** in the IPD (or indeed in any iterated finite game), however, there are strategies **j** different from **i** such that **j** mimics the way **i** plays when it plays against **i** or **j**. The existence of these "neutral mutants" implies that MS cannot be satisfied and so no EPD has usn-stability. This argument, of course, uses the assumption that any strategy in the iterated game is a possible invader. There may be good reason to restrict the available strategies. For example, if the players are assumed to have no knowledge of previous interactions, then it may be appropriate to restrict available strategies to the unconditional ones. Since a pair of players then get the same payoffs in every round of an iterated game, we may as well take each round of the evolutionary game to be one-shot games between every pair of players rather than iterated games. Indeed, this is the kind of evolutionary game that Maynard Smith himself considered. In this framework, any strategy **S** such that $(\mathbf{S}, \mathbf{S})$ is a strict nash equilibrium in the underlying one-shot game (including unconditional defection in the PD) meets the MS condition. Thus MS and usn-stability are non-trivial conditions in some contexts.

A strategy **s** has *restricted weak broad stability)* (rwb-stability) if, when evolution proceeds according to the proportional fitness rule and the native population is playing **s**, any (possibly heterogeneous) group of invaders of sufficiently small size will fail to drive the

natives to extinction. This condition turns out to be equivalent to a weakened version of MS identified by Bendor and Swistak.

> BS) For all strategies $\mathbf{j}$, $V(\mathbf{i}, \mathbf{i}) > V(\mathbf{j}, \mathbf{i})$ or both $V(\mathbf{i}, \mathbf{i})=V(\mathbf{j}, \mathbf{i})$ and $V(\mathbf{i}, \mathbf{j}) \geq V(\mathbf{j}, \mathbf{j})$

BS and rwb-stability are non-trivial conditions in the more general evolutionary framework: strategies for the EPD that satisfy rwb-stability do exist. This does not particularly vindicate any of the strategies discussed above, however. Bendor and Swistak prove a result analogous to the folk theorem mentioned previously: If the shadow of the future is sufficiently large, there are rwb-stable strategies supporting any degree of cooperation from zero to one. One way to distinguish among the strategies that meet BS is by the size of the invasion required to overturn the natives, or, equivalently, by the proportion of natives required to maintain stability. Bendor and Swistak show that this number, the *minimal stabilizing frequency*, never exceeds 1/2: no population can resist every invading group as large as itself. They maintain that this result does allow them to begin to provide a theoretical justification for Axelrod's claims. They are able to show that, as the shadow of the future approaches one, any strategy that is nice (meaning that it is never first to defect) and retaliatory (meaning that it always defects immediately after it has been defected against) has a minimal stabilizing frequency approaching one half. **TFT** has both these properties and, in fact, they are the first two of the four properties Axelrod cited as instrumental to **TFT**'s success. There are, of course, many other nice and retaliatory strategies, and there are strategies (like $\mathbf{P}_1$) that are not retaliatory but still satisfy rwb-stability. But Bendor and Swistak are at least able to show that any "maximally robust" strategy, i.e., any strategy whose minimum stabilizing frequency approaches one half, chooses cooperation on all but a finite number of moves in an infinitely repeated PD.

Bendor and Swistak's results must be interpreted with some care. First, one should keep in mind that no probabilistic or noise- sensitive strategies can fit the definitions of either "nice" or "retaliatory" strategies. Furthermore, imperfect versions of **TFT** do not satisfy rwb-stability. They can be overthrown by arbitrarily small invasions of deterministic **TFT** or, indeed, by arbitrary small invasions of any less imperfect **TFT**. Second, one must remember that the results about minimal stabilizing frequencies only concern weak stability. If the number of generations is large compared with the original population (as it often is in biological applications), a population that is initially composed entirely of players employing the same maximally robust strategy, could well admit a sequence of small invading groups that eventually reduces the original strategy to less than half of the population. At that point the original strategy could be overthrown.

It is likely that both of these caveats play some role in explaining an apparent discrepancy between the Bendor/Swistak results and the Nowak/Sigmund simulations. One would expect Bendor/Swistak's minimal stabilizing frequency to provide some indication of the

length of time that a population plays a particular strategy. A strategy requiring a large invasion to overturn is likely to prevail longer than a strategy requiring only a small invasion. A straightforward calculation reveals that $\mathbf{P}_1$ has a relatively low minimum stabilizing frequency. It is overturned by invasions of unconditional defectors exceeding 10% of the population. Yet in the Nowak/Sigmund simulations, $\mathbf{P}_1$-like strategies predominate over **TFT**-like strategies. Since the simulations required imperfection and since they generated a sequence of mutants vastly larger than the original population, there is no real contradiction here. Nevertheless the discrepancy suggests that we do not yet have a theoretical understanding of EPDs sufficient to predict the strategies that will emerge under various plausible conditions.

Like usn-stability, the concept of rwb-stability can be more discriminating if it is relativized to a particular set of strategies. Molander's 1992 investigation of Schelling's many-person version of the PD, for example, restricts attention to the family $\{\mathbf{S}_1, \ldots, \mathbf{S}_n\}$ of **TFT**-like stratgies. A player adopting $\mathbf{S}_i$ cooperates on the first round and on every subsequent round after at least i others cooperate. By construing stability as resistance to invasions by other family members, Molander is able to show that there are conditions under which a particular mix of two of the $\mathbf{S}_i$'s (one equivalent to **Du**) is uniquely stable. The significance of results like these, however, depends on the plausibility of such limitations on the set of permissible strategies.

17. Spatial PDs

A previous section discussed a controversial argument that cooperation is rational in a PD when each player knows that the other is enough like himself to make it likely that they will choose the same move. An analog of this argument in the evolutionary context is more obviously cogent. If agents are not paired at random, but rather are more likely to play others employing similar strategies, then cooperative behavior is more likely to emerge.

There are at least three mechanisms by which this kind of "association" among players can be achieved. One such mechanism in evolutionary PDs has been widely studied under the label "spatial PD." Players are arranged in some "geographical" arrangement. This may be an array with a rectangular boundary, for example, or a circle, or surface of a sphere or surface of a torus with no boundary. From the geographical arrangement two (possibly identical) kinds of neighborhoods are identified for each player. Agents meet only those in their "interaction" neighborhood, and the evolutionary dynamics considers only the payoffs to those in their "comparison" neighborhood. Generally, the evolutionary dynamics employed is one of "winner imitation" within the interaction neighborhood. (This can model either the idea that each player is invaded by its most successful neighbor or the idea that each player adopts the most successful strategy that it sees.) Because both evolution and interaction are "local," players are more likely (after the first round) to meet those

playing strategies like their own in an SPD than they would be in an ordinary evolutionary game. In addition to the "association" effects, one should also keep in mind that the outcome of SPDs may be influenced by winner imitation dynamics, which may drive to extinction strategies that might survive—and eventually predominate—with the replicator dynamics more commonly employed in ordinary EPDs.

As usual, the impetus for looking at spatial SPDs seems to come from Axelrod. Four copies of each of the 63 strategies submitted to Axelrod's tournament were arranged on a grid with a spherical geometry so that each cell had four neighbors for both interaction and comparison. For every initial random distribution, the resulting SPD eventually reached a state where the strategy in every cell was cooperating with all its neighbors, at which point no further evolution is possible. Only about ten of the 63 original strategies remained in these end-states, and they were no longer randomly distributed, but segregated into clumps of various sizes. Axelrod also showed that under special conditions evolution in an SPD can create successions of complex symmetrical patterns that do not appear to reach any steady-state equilibrium.

To get an idea of why cooperative behavior might spread in this and similar frameworks, consider two agents on either side of a frontier between cooperating and non-cooperating subpopulations. The cooperative agent sees a cooperative neighbor whose four neighbors all cooperate, and who therefore gets four times the reward payoff after playing them all. So he will imitate this neighbor's strategy and remain cooperative. The non-cooperating agent, on the other hand, sees his cooperative counterpart, who gets three reward payoffs from his cooperative neighbors and one sucker payoff. He compares this to the payoffs of his non-cooperatiave neighbors. The best these can do is to get three punishments and a temptation. So, as long as $3R+S$ exceeds $3P+T$, the non-coperative agent on the frontier will adopt the strategy of his cooperative neighbor. Axelrod's payoffs of 5,3,1 and 0 for T, R, P and S, do meet this condition.

Nowak and May have investigated in greater detail SPDs in which the only permitted strategies are **Cu** and **Du**. (These are the strategies appropriate among individuals lacking memory or recognition skills.) They find that, for a variety of spatial configurations, and distributions of strategies evolution depends on relative payoffs in a uniform way. When the temptation payoff is sufficiently high, clusters of **Du** grow and those of **Cu** shrink; when it is sufficiently low, the **Du** clusters shrink and the **Cu** ones grow. For a narrow range of intermediate values, we get more of the complicated patterns noted above. The evolving patterns exhibit great variety. For a given spatial configuration, however, the ratio of the two strategies seems to approach the same constant value for all initial distributions of strategies and all payoffs within the special range. More recently, Nowak, Boenhoeffer and May have observed similar phenomena under a variety of error-conditions identified by Mukherji and Rajan, although cooperators seem to require lower relative temptation

values to thrive under these conditions, and the level of error must be sufficiently low.

Grim, Mar and St Denis report a number of SPD simulations with a greater variety of initial strategies. In general their observations confirm the plausible conjecture that cooperative outcomes are more common in SPDs than ordinary EPDs. Simulations starting with all of the pure reactive strategies of Nowak and Sigmund (i.e., all of the strategies $R(y, p, q)$ described above where y, p and q are either 0 or 1.), all ended with **TFT** (i.e., with $R(1,1,0)$) as the only survivor (though other outcomes—including one in which **Du** is the sole survivor and ones in which **Cu** and **TFT** are intermixed—are clearly possible.) Simulations of the 64 possible pure strategies in which a move may depend on the opponent's previous two moves, ended with mixed populations of survivors all of whom defect after two defections, cooperate after two cooperations and cooperate in the second round of the game. (Again, other outcomes are possible.) Simulations with many (viz., 100) evenly distributed samples of Nowak and Sigmunds mixed reactive strategies, tended to be taken over by $R(.99, .1)$, which is a version of generous **TFT** with slightly less generosity than **GTFT**. Simulations beginning with a random selection of a few (viz., 8) of these strategies tended to evolve to a mixed stable or cyclic pattern dominated by a single version of generous **TFT** with considerably more generosity than **GTFT**. $R(.99, .6)$ seems to have been a frequent victor.

The "geographical" aspect of SPD's need not be taken too literally. For social applications, and probably even for biological ones, there seems to be no motivation for any particular geometrical arrangement. (Why not a "honeycomb," for example, where each agent has six neighbors, rather than a grid where each agent has four?) The interest in SPDs presumably lies in the insight that my "neighborhood" of interaction and my "neighborhood" of comparison, is much smaller than the population as a whole even if it turns out not to be limited by details of physical geography. Nevertheless SPD models of the evolution of cooperation in particular geometrical arrangements have given us some suggestive and pretty pictures to contemplate. Several examples are accessible through the links at the end of this entry.

18. PDs and Social Networks

One way to make the idea of local interaction more realistic for some applications is to let the agents *choose* the partners with whom to interact, based on payoffs in past interactions. Skyrms 2004 considers iterated PDs among a population of unconditional cooperators and defectors. Initially, as usual, each agent chooses a partner at random from the remaining members of the population. For subsequent interactions, however, the odds of choosing that partner are adjusted, according to either the payoffs from previous times that partner was chosen, or (more realistically) the payoffs from previous times that there has been interaction with that partner (regardless of which one was "chooser"). In a typical PD,

where the payoffs for temptation, reward, punishment and sucker are 3,2,1 and 0, both cooperators and defectors eventually choose only cooperators. Since the cooperators are chosen by both cooperators and defectors, they play more often than the defectors who play only when they are chosen. If we assume that there is an equal division between cooperators and defectors, then cooperators can expect a return of two reward payoffs, while defectors can expect a return of one temptation payoff. So with payoff structure indicated, cooperators do better, even with this "one-way" association. The story may unfold somewhat differently in what Skyrms calls an "attenuated" PD, where the payoffs are, let us say, 2.01, 2, 1.98, and 0. (We might think of this as the "Just Don't Be a Sucker" game.) Here, as before, the cooperators quickly learn not to choose defectors as partners. The defectors get roughly the same payoffs whether they choose cooperators or defectors as partners. Since they rapidly cease being chosen by cooperators, however, their returns from interactions with cooperators will be less than returns from defectors and they will soon limit their choices to other defectors. (It is important to understand here that the learning algorithm that determines the probability I will interact with agent **a** depends on *total* returns from interacting with **a** (or total *recent* returns from interacting with **a**) rather than *average* returns from interacting with **a**.) So in the attenuated game we end up with perfect association: defectors play defectors and cooperators play cooperators. Since the reward payoff slightly exceeds the punishment payoff, the cooperators again do better than the defectors.

The social network games considered above are not really evolutionary PDs in the sense described above. The patterns of interaction evolve, but the strategy profile of the population remains fixed. It is natural to allow both strategies and probabilities of interaction to evolve simultaneously as payoffs are distributed. Whether cooperation or defection (or neither) comes to dominate the population under such conditions depends on a multitude of factors: the values of the payoffs, the initial distribution of strategies, the relative speed of the adjustments in strategy and interaction probabilities, and other properties of those two evolutionary dynamics. Skyrms 2004 contains a general discussion and a number of suggestive examples, but it does not provide (or aim to provide) a comprehensive account of social network PDs or a careful analysis of precise formulations to properly model particular phenomena. Much remains unknown.

19. Group Selection and the Haystack PD

A third mechanism by which players can be made more likely to meet those like themselves is to consider a more sophisticated dynamics of evolution that operates on groups of players as well as on the individuals within those groups. There has been a heated debate among biologists and philosophers of biology about the appropriate "units of selection" on which natural selection operates. The idea that in many cases it makes sense to take these units to be groups of individuals (instead of, or in addition to, genes or

individuals) has recently been resuscitated as a respectable and plausible viewpoint. (See Sober and Wilson or Wilson and Sober for a history and impassioned defense of this resuscitation.) For cultural evolution, the idea is equally plausible—within-group behavior may be in equilibrium, but the equilibria reached by different groups may be different. Less successful groups may imitate, be replaced by, or lose members to, more successful groups. Sober and Wilson sometimes write as if evolutionary game theory is an *alternative* viewpoint to group selection, but it is important to understand that this is only true of simple evolutionary models like those presented above. More sophisticated evolutionary games are possible. Consider, for example, a simple version of the *haystack model* originally described by John Maynard Smith. Pairs of players from a large population pair randomly. Each pair colonizes a single haystack. The pair plays a prisoners dilemma and the payoffs to an individual determine the number of offspring of that individual in the next generation. (Parents die when the children are born.) The members of the colony pair randomly with other members and play the PD for some fixed number of generations. Then the haystacks are torn down, the population mixes and random pairs colonize next season's haystacks. One simple way to represent the *n*-generation *haystack PD*, as we might call it is to view it as a game between the two initial founders of a haystack with the payoff to a founder being set to the number of living descendants who are using his strategy. (This idea is suggested in Bergstrom and reported in Skyrms 2004.) For example, suppose $n=3$ and the temptation, reward, punishment and sucker payoffs are set to 5,3,1,0. Then, if Player One cooperates and player Two defects, the payoff to Player One will be 0, because a cooperator gets 0 offspring in the second, and any subsequent, generation. The payoff to player Two will be 5 because the defector has five (like-minded) offspring among the second generation and each of these has one in the third generation since there are no cooperators left to meet. The full payoff matrix for the four generation haystack PD with payoffs 3,2,1, and 0 is given by the matrix below.

	C	D
C	8,8	0,3
D	3,0	1,1

As Skyrms 2004 notes, this matrix characterizes an ordinary Stag Hunt game, as defined above. In fact, Skyrms' observation is generally true. For any PD game *g*, if *n* is sufficiently large, the *n*-generation haystack version of *g* is a Stag Hunt. A simple argument for this result is given in the following very short document:

Haystack PDs Become Stag Hunts.

Bibliography

- Aumann, Robert, "Backward Induction and Common Knowledge of Rationality," *Games and Economic Behavior*, **8** (1995):97-105
- Aumann, Robert, "Note on the Centipede Game," *Games and Economic Behavior*, **23** (1998):97-105
- Axelrod, Robert, "The Emergence of Cooperation Among Egoists," *The American Political Science Review*, **75** (1981): 306-318.
- Axelrod, Robert, *The Evolution of Cooperation* , New York: Basic Books, 1984.
- Axelrod, Robert and Douglas Dion, "The Further Evolution of Cooperation"*Science*, **242**(9 December 1988): 1385-1390.
- Axelrod, Robert and William Hamilton, "The Evolution of Cooperation"*Science*, **211**(27 March 1981): 1390-1396.
- Becker, Neal and Ann Cudd, "Indefinitely Repeated Games: A Response to Carroll," *Theory and Decision*, **28**(1990): 189-195.
- Bendor, Jonathan, "In Good Times and Bad: Reciprocity in an Uncertain World," *American Journal of Political Science*, **31** (1987): 531-558.
- Bendor, Jonathan, "Uncertainty and the Evolution of Cooperation," *Journal of Conflict Resolution*, **37** (1993): 709-733.
- Bendor, Jonathan, and Piotr Swistak, "Types of Evolutionary Stability and the Problem of cooperations," *Proceedings of the National Academy of Sciences*, **92** (April 1995): 3596-3600.
- Bendor, Jonathan, and Piotr Swistak, "The Controversy about the Evolution of Cooperation and the Evolutionary Roots of Social Institutions," in Gasparski, Wojciech et al (eds), *Social Agency*, New Brunswick, N.J.: Transaction Publishers, 1996.
- Bendor, Jonathan, and Piotr Swistak, "Evolutionary Equilibria: Characterization Theorems and Their Implications," *Theory and Decision*, **3** (forthcoming).
- Bendor, Jonathan, and Piotr Swistak, "The Evolutionary Stability of Cooperations," *American Political Science Review*, **91**/2 (June 1997): 290-307.
- Bendor, Jonathan, Roderick Kramer and Piotr Swistak, "Cooperation Under Uncertainty: What is New, What is True and What is Important?" *American Sociological Review*, **61** (April 1996): 333-338.
- Bergstrom, T., "Evolution of Social Behavior: Individual and Group Selection Models," *Journal of Economic Perspectives* **16** (2002): 231-238.
- Bicchieri, Cristina, "Self-refuting Theories of Strategic Interaction," *Erkenntinis*, **30** (1989): 69-85.
- Binmore, Kenneth, *Fun and Games,* Lexington, MA: D.C. Heath and Company (1992).
- Binmore, Kenneth, *Playing Fair: Game Theory and the Social Contract 1*,

Cambridge, MA: MIT Press (1994).

- Binmore, Kenneth, "Rationality and Backward Induction," *Journal of Economic Methodology*, **4** (1997): 23-41.
- Binmore, Kenneth, *Natural Justice* New York, NY: Oxford Univsity Press (2005).
- Boyd, Robert and Jeffrey Lorberbaum, "No Pure Strategy is Evolutionarily Stable in the repeated Prisoner's Dilemma Game," *Nature*, **327** (7 May 1987): 58-59.
- Carroll, J.W. "Indefinite Terminating Points and the Iterated Prisoner's Dilemma," *Theory and Decision* **22** (1987):247-256.
- Cambell, Richmond and Lanning Snowden, *Paradoxes of Rationality and Cooperation*, Vancouver: University of British Columbia Press (1985).
- Danielson, Peter, *Artificial Morality: Virtual Robots for Virtual Games*, London: Routledge (1992).
- Davis, Laurence, "Prisoners, Paradox and Rationality," *Amierican Philosophical Quarterly*, **14** (1977): 319-327. Reprinted in Campbell and Snowden, 45-58.
- Davis, Laurence, "Is the Symmetry Argument Valid?," in Campbell and Snowden (1985), 255-262.
- Donninger, Christian "Is It Always Efficient to be Nice?" in Dickman and Mitter (eds) *Paradoxical Effects of Social Behavior* Heidelberg: Physica Verlag (1986): 123-134.
- Farrell, Joseph, and Roger Ware, "Evolutionary Stability in the Repeated Prisoner's Dilemma," *Theoretical Population Biology*, **36** (1989): 161-167.
- Gauthier, David, *Morals by Agreement*, Oxford: Clarendon Press (1986).
- Grim, Patrick, Gary Mar and Paul St. Denis, *The Philosophical Computer*, Cambrige, Mass: MIT Press (1998).
- Hardin, Garret, "The Tragedy of the Commons," *Science*, **162** (13 December 1968): 1243-1248.
- Howard, Nigel, *Paradoxes of Rationality*, Cambridge, MA: MIT Press (1971).
- Howard, J.V. "Cooperation in the Prisoner's Dilemma," *Theory and Decision*, **24** (1988):203-213.
- Hurley, S.L. "Newcomb's Problem, Prisoners' Dilemma, and Collective Action," *Synthese*, **86** (1991):173-196.
- Joyce, James, *The Foundations of Causal Decision Theory*, Cambridge University Press, (1999).
- Kavka, Gregory, "Hobbes War of All Against All," *Ethics*, **93** (1983): 291-310.
- Kavka, Gregory, *Hobbesean Moral and Political Theory*, Princeton: Princeton University Press (1986).
- Kavka, Gregory, "Is Individual Choice Less Problematic than Collective Choice?" *Economics and Philosophy*, **7** (1991): 291-310.
- Kollock, Peter, "An Eye For an Eye Leaves Everybody Blind: Cooperation and Accounting Systems," *American Sociological Review*, **58** (1993): 768-786.

- Kraines, David and Vivian Kraines, "Pavlov and the Prisoner's Dilemma," *Theory and Decision*, **26** (1989): 47-79.
- Kraines, David and Vivian Kraines, "Learning to Cooperate with Pavlov: an Adaptive Strategy for the Iterated Prisoner's Dilemma with Noise," *Theory and Decision*, **35** (1993): 107-150.
- Kreps, David, Paul Milgrom, John Roberts and Robert Wilson, "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma," *Journal of Economic Theory*, **27** (1982): 245-252.
- Kuhn, Steven, "Agreement Keeping and Indirect Moral Theory" *Journal of Philosophy*, **93** (1996): 105-128.
- Kuhn, Steven, "Reflections on Ethics and Game Theory" *Synthese*, **141** (2004): 1-44.
- Kuhn, Steven, and Serge Moresi, "Pure and Utilitarian Prisoner's Dilemmas" *Economics and Philosophy*, **11** (1995): 123-133.
- Lewis, David, "Prisoner's Dilemma Is a Newcomb Problem," *Philosophy and Public Affairs* **8** (1979): 235-240.
- Maynard Smith, John, "The Evolution of Behavior," *Scientific American*, **239** (1978): 176-192.
- Molander, Per, "The Optimal Level of Generosity in a Selfish, Uncertain Environment," *Journal of Conflict Resolution*, **29**, (December 1985): 611-619.
- Molander, Per, "The Prevalence of Free Riding," *Journal of Conflict Resolution*, **36**, (December 1992): 756-771.
- Mukherji, Arijit, Vijay Rajan and James Slagle, "Robustness of Cooperation," *Nature*, **379** (11 January 1996): 125-126.
- Nowak, Martin, and Robert May, "Evolutionary Games and Spatial Chaos,"*Nature*, **359** (29 October 1992): 826-829.
- Nowak, Martin and Karl Sigmund, "Tit for Tat in Heterogeneous Populations," *Nature*, **355** (16 January 1992): 250-253.
- Nowak, Martin and Karl Sigmund, "A Strategy of Win-stay, Lose-shift that Outperforms Tit-for-tat in the Prisoner's Dilemma Game," *Nature*, **364** (1 July 1993): 56-58.
- Nowak, Martin, Robert May, and Karl Sigmund, "The Arithmetics of Mutual Help," *Scientific American* (June 1995): 76-81.
- Nozick, Robert, "Newcomb's Problem and Two Principles of Choice", in N. Resher (ed) *Essays in Honor of Carl G. Hempel*, Reidel (1969):114-146. Reprinted in Cambell and Snowden, 107-132.
- Pettit, Phillip, "Free Riding and Foul Dealing," *Journal of Philosophy*, **83** 1986: 361-379.
- Pettit, Phillip and Robert Sugden, "The Backward Induction Paradox," *Journal of Philosophy*, **86** 1989: 169-182.
- Poundstone, William, *Prisoner's Dilemma* New York: Doubleday (1992).

- Quinn, Warren, "The Paradox of the Self-Torturer," *Philosophical Studies*, **59** (1990):79-90.
- Rabinowicz, Wlodek, "Grappling with the Centipede: Defense of Backward Induction for BI-Terminating Games," *Economics and Philosophy*,, **14** (1998): 95-126.
- Rosenthal, R. "Games of Perfect Information, Predatory Pricing, and the Chain Store," *Journal of Economic Theory* **25** (1981): 92-100.
- Schelling, Thomas, *Micromotives and Macrobehavior* New York: Norton (1978).
- Segal, Nancy and Scott Hershberger,"Cooperation and Competition Between Twins: Findings from a Prisoner's Dilemma Game," *Evolution and Human Behavior* **20**, 1999:29-51
- Selten, Reinhard, "Reexamination of the Perfectness Concept of Equilibrium in Extensive Games," *International Journal of Game Theory*,, **4** (1975): 25-55.
- Selten, Reinhard, "The Chain-Store Paradox," *Theory and Decision*, **9** (1978): 127-159.
- Selten, Reinhard, "Evolutionary Stability in Extensive Two-person Games," *Mathematical Social Sciences*, **5** (1983): 269-363.
- Sigmund, Karl, *Games of Life: Explorations in Ecology Evolution and Behavior*, Oxford: Oxford University Press (1993).
- Skyrms, Brian, *The Dynamics of Rational Deliberation*, Cambridge, MA: Harvard University Press (1990).
- Skyrms, Brian, *Evolution of the Social Contract*, Cambridge, Cambridge University Press (1996).
- Skyrms, Brian, "The Shadow of the Future," in Coleman and Morris (eds.)*Rational Commitment and Social Justice: Essays for Gregory Kavka*, New York, Cambridge University Press (1998).
- Skyrms, Brian, *The Stag Hunt and the Evolution of Social Structure*, Cambridge, Cambridge University Press (2004).
- Sobel, J.H., "Backward Induction Without Tears?," in D. Vanderveken (ed), *Logic, Thought and Action*,, Springer (2005):433-461.
- Sobel, J.H., "Backward Induction Arguments: A Paradox Regained," *Philosophy of Science* **60** 1993: 114-133.
- Sober, Elliott and David Sloan Wilson, *Unto Others: The Evolution and Psychology of Unselfish Behavior*, Cambridge, MA: 1998.
- Sugden *The Economics of Rights, Cooperation and Welfare*, New York, Basil Blackwell (1986). Second edition published in 2004 by Palgrave MacMillan, Basingstoke, UK.
- Taylor, Michael, *The Possibility of Cooperation*, Cambridge: Cambridge University Press (1987).
- Trivers, Robert, "The Evolution of Reciprocal Altruism," *Quarterly Review of*

Biology, **46** (1971): 35-57.

- Vanderschraaf, Peter, "The Informal Game Theory in Hume's Account of Convention," *Economics and Philosophy,* **14**(1998):215-247.
- Williamson, Timothy, "Inexact Knowledge," *Mind*, **101** (1992): 217-242.
- Wilson, D.S. and E. Sober, "Reintroducing Group Selection to the Human Behavioral Sciences," *Behavioral and Brain Sciences*, **17** (1994): 585-654.

Other Internet Resources

- "Comprehsive Repository" for information about the iterated PD (compiled by group at Laboratoire d'Informatique Fondamentale de Lille).
- Spatial IPDs by Norman Siebrasse
- Miscellaneous PD Resources (compiled by Constitution Society).
- Interactive Prisoner's Dilemma (at the Serendip pages at Bryn Mawr).

Related Entries

game theory | game theory: and ethics | game theory: evolutionary