

The Prisoner's Dilemma Computer Tournaments and the Evolution of Cooperation

May, 1983

LIFE is filled with paradoxes and dilemmas. Sometimes it even feels as if the essence of living is the sensing-indeed, the savoring-of paradox. Although all paradoxes seem somehow related, some paradoxes seem abstract and philosophical, while others touch on life very directly. A very lifelike paradox is the so-called "Prisoner's Dilemma", discovered in 1950 by Melvin Dresher and Merrill Flood of the RAND Corporation. Albert W. Tucker wrote the first article on it, and in that article he gave it its now-famous name. I shall here present the Prisoner's Dilemma-first as a metaphor, then as a formal problem.

The original formulation in terms of prisoners is a little less clear to the uninitiated, in my experience, than the following one. Assume you possess copious quantities of some item (money, for example), and wish to obtain some amount of another item (perhaps stamps, groceries, diamonds). You arrange a mutually agreeable trade with the only dealer of that item known to you. You are both satisfied with the amounts you will be giving and getting. For some reason, though, your trade must take place in secret. Each of you agrees to leave a bag at a designated place in the forest, and to pick up the other's bag at the other's designated place. Suppose it is clear to both of you that the two of you will never meet or have further dealings with each other again.

Clearly, there is something for each of you to fear: namely, that the other one will leave an empty bag. Obviously, if you both leave full bags, you will both be satisfied; but equally obviously, getting something for clothing is even more satisfying. So you are tempted to leave an empty bag. In fact, you can even reason it through quite rigorously this way: "If the dealer brings a full bag, I'll be better off having left an empty bag, because I'll have gotten

all that I wanted and given away nothing. If the dealer brings an empty bag, I'll be better off having left an empty bag, because I'll not have been cheated I'll have gained nothing but lost nothing either. Thus it seems that no matter what the dealer chooses to do, I'm better off leaving an empty bag. So I'll leave an empty bag."

The dealer, meanwhile, being in more or less the same boat (though at the other end of it), thinks analogous thoughts and comes to the parallel conclusion that it is best to leave an empty bag. And so both of you, with your impeccable (or impeccable-seeming) logic, leave empty bags, and go away empty-handed. How sad, for if you had both just cooperated, you could have each gained something you wanted to have. Does logic prevent cooperation? This is the issue of the Prisoner's Dilemma.

* * *

In case you're wondering why it is called "Prisoner's Dilemma", here's the reason. Imagine that you and an accomplice (someone you have no feelings for one way or the other) committed a crime, and now you've both been apprehended and thrown in jail, and are fearfully awaiting trials. You are being held in separate cells with no way to communicate. The prosecutor offers each of you the following deal (and informs you both that the identical deal is being offered to each of you-and that you both know that as well): "We have a lot of circumstantial evidence on you both. So if you both claim innocence, we will convict you anyway and you'll both get two years in jail. But if you will help us out by admitting your guilt and making it easier for us to convict your accomplice-oh, pardon me, your *alleged* accomplice why, then, we'll let you out free. And don't worry- about revenge-your accomplice will be in for five years! How about it?" Warily you ask, "But what if we *both* say we're guilty?" "Ah, well, my, friend-I'm afraid you'll both get four-year sentences, then."

Now you're in a pickle! Clearly, you don't want to claim innocence if your partner has sung, for then you're in for five long years. Better you should both have sung-then you'll only get four. On the other hand, if your partner claims innocence, then the best possible thing for you to do is sing. since then you're out scot-free! So at first sight, it seems obvious what you should do: Sing! But what is obvious to you is equally obvious to your opposite number, so now it looks like you both ought to sing, which means -Sing Sing for four years! At least that's what logic tells you to do. Funny% since if both of you had just been illogical and maintained innocence, you'd both be in for only half as long! Ah, logic does it again.

* * *

Let us now go back to the original metaphor and slightly alter its conditions. Suppose that both you and your partner very much want to have

a regular supply of what the other has to offer, and so, before conducting your first exchange, you agree to carry on a lifelong exchange, once a month. You still expect never to meet face to face. In fact, neither of you has any idea how old the other one is, so you can't be very sure of how long this lifelong agreement may go on, but it seems safe to assume it'll go on for a few months anyway, and very likely for years.

Now, what do you do on your first exchange? Taking an empty bag seems fairly nasty as the opening of a relationship—hardly an effective way to build up trust. So suppose you take a full bag, and the dealer brings one as well. Bliss—for a month. Then you both must go back. Empty, or full? Each month, you have to decide whether to *defect* (take an empty bag) or to *cooperate* (take a full one). Suppose that one month, unexpectedly, your dealer defects. Now what do you do? Will you suddenly decide that the dealer can never be trusted again, and from now on always bring empty bags, in effect totally giving up on the whole project forever? Or will you pretend you didn't notice, and continue being friendly? Or will you try to punish the dealer by some number of defections of your own? One? Two? A random number? An increasing number, depending on how many defections you have experienced? Just how mad will you get?

This is the so-called *iterated* Prisoner's Dilemma. It is a very difficult problem. It can be, and has been, rendered more quantitative and in that form studied with the methods of game theory and computer simulation. How does one quantify it? One builds a *payoff matrix* presenting point values for the various alternatives. A typical one is shown in Figure 29-1a. In this matrix, mutual cooperation earns both parties 2 points (the subjective value of receiving a full bag of what you need while giving up a full bag of what you have). Mutual defection earns you both 0 points (the subjective value of gaining nothing and losing nothing, aside from making a vain trip out to the forest that month). Cooperating while the other defects stings: you get -1 point while the rat gets 4 points! Why so many? Because it is so pleasurable to get something for nothing. And of course, should you happen to be a rat some month when the dealer has cooperated, then you get 4 points and the dealer loses 1.

It is obvious that in a *collective* sense, it would be best for both of you to always cooperate. But suppose you have no regard whatsoever for the other. There is no "collective good" you are both working for. You are both supreme egoists. Then what? The meaning of this term, "egoist", can perhaps be made clear by the following. Suppose you and your dealer have developed a trusting relationship of Mutual cooperation over the years, when one day you receive secret and reliable information that the dealer is quite sick and will soon die, probably within a month or two. The dealer has no reason to suspect that you have heard this. Aren't you highly tempted to defect, all of a sudden, despite all your years of cooperating? You are, after all, out for yourself and no one else in this cruel, cruel world. And since it seems that this may very well be the dealer's last month, why not profit

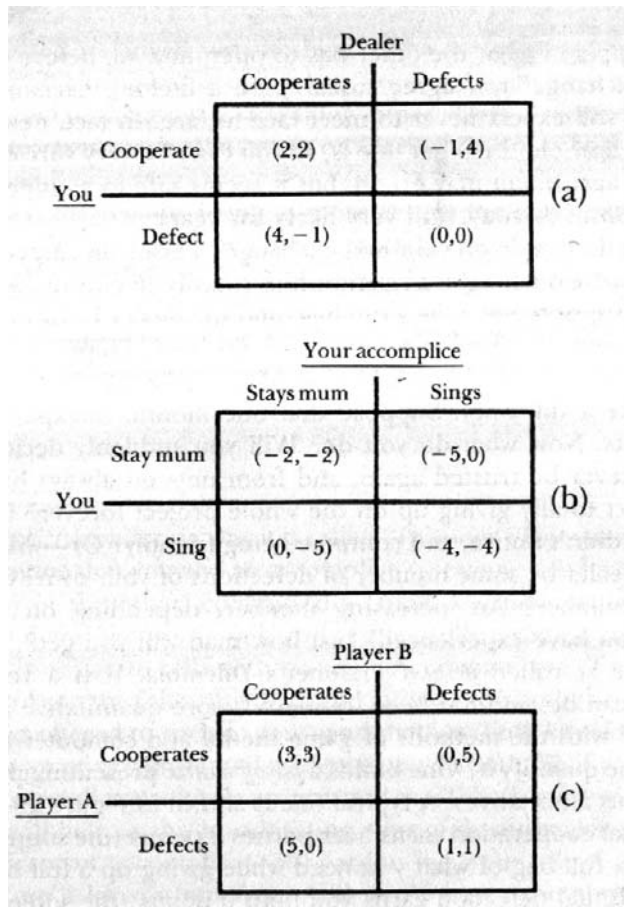


FIGURE 29-1. The Prisoner's Dilemma.

In (a), a Prisoner's Dilemma payoff matrix in the case of a dealer and a buyer of commodities or services, in which both participants have a choice: to cooperate (i.e., to deliver the goods or the payment) or to defect (i.e., to deliver nothing). The numbers attempt to represent the degree of satisfaction of each partner in the transaction.

In (b), the formulation of the Prisoner's Dilemma to which it owes its name: in terms of prisoners and their opportunities for double-crossing or collusion. The numbers are negative because they represent punishments: the length of both prisoners' prospective jail sentences, in years. This metaphor is due to Albert W Tucker.

*In (c), a Prisoner's Dilemma formulation where all payoffs are nonnegative numbers. This is my canonical version, following the usage in Robert Axelrod's book, *The Evolution of Cooperation*.*

as much as possible from your secret knowledge? Your defection may never be punished, and at the worst, it will be punished by one last-gasp defection by the dying dealer.

The surer you are that this next turn is to be the very last one, the more you feel you must defect. Either of you would feel that way, of course, on learning that the other one was nearing the end of the rope. This is what is meant by "egoism". It means you have no feeling of friendliness or

goodwill or compassion for the other player; you have no conscience; all you care about is amassing points, more and more and more of them.

What does the payoff matrix for the other metaphor, the one involving prisoners, look like? It is shown in Figure 29-1b. The equivalence of this matrix to the previous matrix is clear if you add a constant—namely, 4—to all terms in this one. Indeed, we could add any constant to either matrix and the dilemma would remain essentially unchanged. So let us add 5 to this one so as to get rid of all negative payoffs. We get the canonical Prisoner's Dilemma payoff matrix, shown in Figure 29-1c. The number 3 is called the reward for mutual cooperation, or R for short. The number 1 is called the punishment, or P. The number 5 is T, the temptation, and 0 is S, the sucker's payoff. The two conditions that make a matrix represent a Prisoner's Dilemma situation are these:

$$(1) T > R > P > S$$

$$(2) (T+S)/2 < R$$

The first one simply makes the argument go through for each of you, that "it is better for me to defect no matter what my counterpart does". The second one simply guarantees that if you two somehow get locked into out-of-phase alternations (that is, "you cooperate, I defect" one month and "you defect, I cooperate" the next), you will not do better—in fact, you will do worse—than if you were cooperating with each other each month.

Well, what would be your best strategy? It can be shown quite easily that there is no universal answer to this question. That is, there is no strategy that is better than all other strategies under all circumstances. For consider the case where the other player is playing ALL D—the strategy of defecting each round. In that case, the best you can possibly do is to defect each time as well, including the first. On the other hand, suppose the other player is using the Massive Retaliatory Strike strategy, which means "I'll cooperate until you defect and thereafter I'll defect forever." Now if you defect on the very first move, then you'll get one T and all P's thereafter until one of you dies. But if you had waited to defect, you could have benefited from a relationship of mutual cooperation, amassing many R's beforehand. Clearly that bunch of R's will add up to more than the single T if the game goes on for more than a few moves. This means that against the ALL D strategy, ALL D is the best counterstrategy, whereas "Always cooperate unless you learn that you or the other player is just about to die, in which case defect" is the best counterstrategy against Massive Retaliatory Strike. This simple argument shows that how you should play depends on who you're playing.

The whole concept of the "quality" of a strategy takes on a decidedly more operational and empirical meaning if one imagines an ocean populated by dozens of little beings swimming around and playing Prisoner's Dilemma over and over with each other. Suppose that each time two such beings encounter each other, they recognize each other and

remember how previous encounters have gone. This enables each one to decide what it wishes to do this time. Now if each organism is continually swimming around and bumping into the others, eventually, each one will have met even other one numerous times, and thus all strategies will have been given the opportunity to interact with each other. By "interact", what is meant here is certainly not that anyone knocks anyone else out of the ocean, as in an elimination tournament. The idea is simply that each organism gains zero or more points in each meeting, and if sufficient time is allowed to elapse, everybody will have met with everybody else about the same number of times. and now the only question is: Who has amassed the most points? Amassing points is truly the name of the game.

It doesn't matter if you have "beaten" anyone, in the sense of having gained more points interacting with them than they gained from interacting with you. That kind of "victory" is totally irrelevant here. What matters is not the number of "victories" rung up by any individual. but the individual's total point count—a number that measures the individual's overall viability in this particular "sea" of many strategies. It sounds nearly paradoxical. but you could lose many—indeed, all—of your individual skirmishes, and yet still come out the overall winner.

As the image suggests very strongly, this whole situation is highly relevant to questions in evolutionary biology. Can totally selfish and unconscious organisms living in a common environment come to evolve reliable cooperative strategies? Can cooperation emerge in a world of pure egoists? In a nutshell, can cooperation evolve out of noncooperation? If so, this has revolutionary import for the theory of evolution, for many of its critics have claimed that this was one place that it was hopelessly snagged.

Well, as it happens, it has now been demonstrated rigorously and definitively that such cooperation can emerge, and it was done through a computer tournament conducted by political scientist Robert Axelrod of the Political Science Department and the Institute for Public Policy Studies of the University of Michigan in Ann Arbor. More accurately, Axelrod first studied the ways that cooperation evolved by means of a computer tournament, and when general trends emerged, he was able to spot the underlying principles and prove theorems that established the facts and conditions of cooperation's rise from nowhere. Axelrod has written a fascinating and remarkably thought-provoking book on his findings, called *The Evolution of Cooperation*, published in 1984 by Basic Books, Inc. (Quoted sections below are taken from an early draft of that book.) Furthermore, he and evolutionary biologist William D. Hamilton have worked out and published many of the implications of these discoveries for evolutionary theory. Their work has won much acclaim—including the 1981 Newcomb Cleveland prize, a prize awarded annually by the American Association for the Advancement of Science for "an outstanding paper published in *Science*".

There are really three aspects of the question "Can cooperation emerge in a world of egoists?" The first is: How can it get started at all? the second is: Can cooperative strategies survive better than their non-cooperative rivals? The third one is: Which cooperative strategies will do the best, and how will they come to predominate?

To make these issues vivid, let me describe Axelrod's tournament and its somewhat astonishing results. In 1979, Axelrod sent out invitations to a number of professional game theorists, including people who had published articles on the Prisoner's Dilemma. telling them that he wished to pit many strategies against one another in a round-robin Prisoner's Dilemma tournament, with the overall goal being to amass as many points as possible. He asked for strategies to be encoded as computer programs that could respond to the 'C' or 'D' of another player, taking into account the remembered history of previous interactions with that same player. A program should always reply with a 'C' or a 'D', of course, but its choice need not be deterministic. That is, consultation of a random-number generator was allowed at any point in a strategy.

Fourteen entries were submitted to Axelrod, and he introduced into the field one more program called RANDOM, which in effect flipped a coin (computationally simulated, to be sure) each move, cooperating if heads came up, defecting otherwise. The field was a rather variegated one, consisting of programs ranging from as few as four lines to as many as 77 lines (of Basic). Every program was made to engage each other program (and a clone of itself) 200 times. No program was penalized for running slowly. The tournament was actually run five times in a row, so that ,pseudo-effects caused by statistical fluctuations in the random-number generator would be smoothed out by averaging.

The program that won was submitted by the old Prisoner's Dilemma hand, Anatol Rapoport, a psychologist and philosopher from the University of Toronto. His was the shortest of all submitted programs, and is called TIT FOR TAT'.. TIT FOR TAT uses a very simple tactic:

Cooperate on move 1;
Thereafter, do whatever the other player did the previous move.

That is all. It sounds outrageously simple. How in the world could such a program defeat the complex stratagems devised by other experts?

Well, Axelrod claims that the game theorists in general did not go far enough in their analysis. They looked "only two levels deep", when in fact they should have looked three levels deep to do better. What precisely does this mean? He takes a specific case to illustrate his point. Consider the entry called JOSS (submitted by Johann Joss, a mathematician from Zurich, Switzerland). JOSS's strategy is very similar to TIT FOR TAT's, in that it

begins by cooperating, always responds to defection by defecting and nearly always responds to cooperation by cooperating. The hitch is that JOSS uses a random-number generator to help it decide when to pull a "surprise defection" on the other player. JOSS is set up so that it has a 10 percent probability of defecting right after the other player has cooperated.

In playing TIT FOR TAT, JOSS will do fine until it tries to catch TIT FOR TIT off guard. When it defects, TIT FOR TAT retaliates with a single defection, while JOSS "innocently" goes back to cooperating. Thus we have a "DC" pair. On the next move, the 'C' and 'D' will switch places since each program in essence echoes the other's latest move, and so it will go: CD, then DC, CD, DC, and so on. There may ensue a long reverberation set off by JOSS's D, but sooner or later, JOSS will randomly, throw in another unexpected D after a C from TIT FOR TAT. At this point, there will be a "DD" pair, and that determines the entire rest of the match. Both will defect forever, now. The "echo" effect resulting from JOSS's first attempt at exploitation and TIT FOR TAT's simple punitive act lead ultimately to complete distrust and lack of cooperation.

This may seem to imply that both strategies are at fault and will suffer for it at the hands of others, but in fact the one that suffers from it most is JOSS, since JOSS tries out the same trick on partner after partner, and in many cases this leads to the same type of breakdown of trust, whereas TIT FOR TAT, never defecting first, will never be the initial cause of a breakdown of trust. Axelrod's technical term for a strategy that never defects before its opponent does is nice. TIT FOR TAT is a nice strategy, JOSS is not. Note that "nice" does not mean that a strategy never defects! TIT FOR TAT defects when provoked, but that is still considered being "nice".

Axelrod summarizes the first tournament this way:

A major lesson of this tournament is the importance of minimizing echo effects in an environment of mutual power. A sophisticated analysis must go at least three levels deep. First is the direct effect of a choice. This is easy, since a defection always earns more than a cooperation. Second are the indirect effects, taking into account that the other side may or may not punish a defection. This much was certainly appreciated by many of the entrants. But third is the fact that in responding to the defections of the other side, one may be repeating or even amplifying one's own previous exploitative choice. Thus a single defection may be successful when analyzed for its direct effects, and perhaps even when its secondary effects are taken into account. But the real costs may be in the tertiary effects when one's own isolated defections turn into unending mutual recriminations. Without their realizing it, many of these rules actually wound up punishing themselves. With the other player serving as a mechanism to delay the self-punishment by a few moves, this aspect of self-punishment was not perceived by the decision rules

The analysis of the tournament results indicates that there is a lot to be learned about coping in an environment of mutual power. Even expert strategists from political science, sociology, economics, psychology, and mathematics made the systematic errors of being too competitive for their own

good. not forgiving enough, and too pessimistic about the responsiveness of the other side.

Axelrod not only analyzed the first tournament, he even performed a number of "subjunctive replays" of it, that is, replays with different sets of entries. He found, for instance, that the strategy called TIT FOR TWO T.ATS, which tolerates two defections before getting mad (but still only- strikes back once), would have won, had it been in the line-up. Likewise, two other strategies he discovered, one called REVISED DOWNING and one called LOOK-AHEAD, would have come in first had they been in the tournament.

In summary, the lesson of the first tournament seems to have been that it is important to be nice ("don't be the first to defect") and forgiving, ("don't hold a grudge once you've vented your anger"). TIT FOR TAT possesses both these qualities, quite obviously.

* * *

After this careful analysis, Axelrod felt that significant lessons had been unearthed, and he felt convinced that more sophisticated strategies could be concocted, based on the new information. Therefore he decided to hold a second, larger computer tournament. For this tournament, he not only invited all the participants in the first round, but also advertised in computer hobbyist magazines, hoping to attract people who were addicted to programming and who would be willing to devote a good deal of time to working out and perfecting their strategies. To each person who entered, Axelrod sent a full and detailed analysis of the first tournament, along with a discussion of the "subjunctive replays" and the strategies that would have won. He described the strategic concepts of "niceness" and "forgiveness" that seemed to capture the lessons of the tournament so well, as well as strategic pitfalls to avoid. Naturally, each entrant realized that all the other entrants had received the same mailing, so that everyone knew that everyone knew that everyone knew that ...

There was a large response to Axelrod's call for entries. Entries were received from six countries, from people of all ages, and from eight different academic disciplines. Anatol Rapoport entered again, resubmitting TIT FOR TAT (and was the only one to do so, even though it was explicitly stated that anyone could enter any program written by anybody). A ten-year-old entered, as did one of the world's experts on game theory and evolution, John Maynard Smith, professor of biology at the University of Sussex in England, who submitted TIT FOR TWO TATS. Two people separately submitted REVISED DOWNING.

Altogether, 62 entries were received, and generally speaking, they were of a considerably higher degree of sophistication than those in the first tournament. The shortest was again TIT FOR TAT, and the longest was a program from New Zealand, consisting of 152 lines of Fortran. Once again,

RANDOM was added to the field, and with a flourish and a final carriage return, the horses were off. Several hours of computer time later, the results came in.

The outcome was nothing short of stunning: TIT FOR TAT, the simplest program submitted, won again. What's more, the two programs submitted that had won the subjunctive replays of the first tournament now turned up way down in the list: TIT FOR TWO TATS came in 24th, and REVISED DOWNING ended up buried in the bottom half of the field.

This may seem horribly nonintuitive, but remember that a program's success depends entirely on the environment in which it is swimming. There is no single "best strategy" for all environments, so that winning in one tournament is no guarantee of success in another. TIT FOR TAT has the advantage of being able to "get along well" with a great variety of strategies, while other programs are more limited in their ability to evoke cooperation. Axelrod puts it this way:

What seems to have happened is an interesting interaction between people who drew one lesson and people who drew another lesson from the first round. Lesson One was "Be nice and forgiving." Lesson Two was more exploitative: "If others are going to be nice and forgiving, it pays to try to take advantage of them." The people who drew Lesson One suffered in the second round from those who drew Lesson Two.

Thus the majority of participants in the second tournament really had not grasped the central lesson of the first tournament: the importance of being willing to initiate and reciprocate cooperation. Axelrod feels so strongly about this that he is reluctant to call two strategies playing against each other "opponents"; in his book he always uses neutral terms such as "strategies" or "players". He even does not like saying they are playing against each other, preferring "with". In this article, I have tried to follow his usage, with occasional departures. One very striking fact about the second tournament is the success of "nice" rules: of the top fifteen finishers, only one (placing eighth) was not nice. Amusingly, a sort of mirror image held: of the bottom fifteen finishers, only one was nice!

Several non-nice strategies featured rather tricky probes of the opponent (sorry!), sounding it out to see how much it "minded" being defected against. Although this kind of probing by a program might fool occasional opponents, more often than not it backfired, causing severe breakdowns of trust. Altogether, it turned out to be very costly to try to use defections to "flush out" the other player's weak spots. It turned out to be more profitable to have a policy of cooperation as often as possible, together with a willingness to retaliate swiftly against any attempted undercutting. Note, however, that strategies featuring massive retaliation were less successful than TIT FOR TAT with its more gentle policy of restrained retaliation.

Forgiveness is the key here, for it helps to restore the proverbial "atmosphere of mutual cooperation" (to use the phrase of international diplomacy) after a small skirmish.

"Be nice and forgiving" was in essence the overall lesson of the first tournament. Apparently, though, many people just couldn't get themselves to believe it, and were convinced that with cleverer trickery and scheming, they could win the day. It took the second tournament to prove them dead wrong. And out of the second tournament, a third key strategic concept emerged: that of provocability-the notion that one should "get mad" quickly at defectors. and retaliate. Thus a more general lesson is: "Be nice, provokable, and forgiving."

* * *

Strategies that do well in a wide variety of environments are called by Axelrod robust. and it seems that ones with "good personality traits"-that is, nice, provokable, and forgiving strategies-are sure to be robust. TIT FOR TAT is by no means the only possible strategy with these traits, but it is the canonical example of such a strategy, and it is astonishingly robust.

Perhaps the most vivid demonstrations of TIT FOR TAT's robustness were provided by various subjunctive replays of the second tournament. The principle behind any replay involving a different environment is quite simple. From the actual playing of the tournament, you have a 63X63 matrix documenting how well each program did against each other program. Now, the effective "population" of a program in the environment can be manipulated mathematically by attaching a weight factor to all that program's interactions, then just retotaling all the columns. This way you can get subjunctive instant replays without having to rerun the tournament.

This simple observation means that the results of a huge number of potential subjunctive tournaments are concealed in, but potentially extractable from, the 63X63 matrix of program-z s.-program totals. For instance, Axelrod discovered, using statistical analysis, that there were essentially six classes of strategies in the second tournament. For each of these classes, he conducted a subjunctive instant replay of the, tournament by quintupling the importance (the weight factor) of that class alone, thus artificially inflating a certain strategic style's population in the environment. When the scores were retotaled, TIT FOR TAT emerged victorious in five out of six of those hypothetical tournaments, and in the sixth it placed second.

Undoubtedly the most significant and ingenious type of subjunctive replay that Axelrod tried was the ecological tournament. Such a tournament consists not merely of a single subjunctive replay, but of -a whole cascade of hypothetical replays, each one's environment determined by the results of the previous one. In particular, if you take a program's score in a tournament as a measure of its "fitness", and if you interpret "fitness" as meaning "number of progeny in the next generation", and finally, if you let

"next generation" mean "next tournament", then what you get is that each tournament's results determine the environment of the next one- and in particular, successful programs become more copious in the next tournament. This type of iterated tournament is called "ecological" because it simulates ecological adaptation (the shifting of a fixed set of species' populations according to their mutually defined and dynamically developing environment), as contrasted with evolution via mutation (where new species can come into existence).

As one carries an ecological tournament through generation after generation, the environment gradually changes. In a paraphrase of how Axelrod puts it, here is what happens. At the very beginning, poor programs and good programs alike are equally represented. As time passes, the poorer ones begin to drop out while the good ones flourish. But the rank order of the good ones may now change, because their "goodness" is no longer being measured against the same field of competitors as initially. Thus success breeds ever more success-but only provided that the success derives from interaction with other similarly successful programs. If, by contrast, some program's success is due mostly to its ability to milk "dumber" programs for all they're worth, then as those programs are gradually squeezed out of the picture, the exploiter's base of support will be eroded and it will suffer a similar fate.

A concrete example of ecological extinction is provided by HARRINGTON, the only non-nice program among the top fifteen finishers in the second tournament. In the first 200 generations of the ecological tournament, while TIT FOR TAT and other successful nice programs were gradually increasing their percentage of the population, HARRINGTON ' too was increasing its percentage. This was a direct result of HARRINGTON'S exploitative strategy. However, by the 200th generation, things began to take a noticeable turn. Weaker programs were beginning to go extinct, which meant fewer and fewer dupes for HARRINGTON to profit from. Soon the trend became apparent: HARRINGTON could not keep up with its nice rivals. By the 1,000th generation, HARRINGTON was as extinct as the dodos it had exploited. Axelrod summarizes:

Doing well with rules that do not score well themselves is eventually a self-defeating process. Not being nice may look promising at first, but in the long run it can destroy the very environment it needs for its own success.

Needless to say, TIT FOR TAT fared spectacularly well in the ecological tournament, increasing its lead ever more. After 1,000 generations, not only was TIT FOR TAT ahead, but its rate of growth was greater than that of any other program. This is an almost unbelievable success story,, all the more so because of the absurd simplicity of the "hero". One amusing aspect of it is that TIT FOR TAT did not defeat a single one of its rivals in their encounters. This is not a quirk; it is in the nature of TIT FOR TAT. TIT FOR TAT simply cannot defeat anyone; the best it can achieve is a tie, and often it loses (though not by much).

Axelrod makes this point very clear:

TIT FOR TAT won the tournament, not by beating the other player, but by eliciting behavior from the other player which allowed both to do well. TIT FOR TAT was so consistent at eliciting mutually rewarding outcomes that it attained a higher overall score than any other strategy in the tournament.

So in a non-zero-sum world you do not have to do better than the other player to do well for yourself. This is especially true when you are interacting with many different players. Letting each of them do the same or a little better than you is fine, as long as you tend to do well yourself. There is no point in being envious of the success of the other player, since in an iterated Prisoner's Dilemma of long duration the other's success is virtually a prerequisite of your doing well for yourself.

He gives examples from everyday life in which this principle holds. Here is one:

A firm that buys from a supplier can expect that a successful relationship will earn profit for the supplier as well as the buyer. There is no point in being envious of the supplier's profit. Any attempt to reduce it through an uncooperative practice, such as by not paying your bills on time, will only encourage the supplier to take retaliatory action. Retaliatory action could take many forms, often without being explicitly labeled as punishment. It could be less prompt deliveries, lower quality control, less forthcoming attitudes on volume discounts, or less timely news of anticipated market conditions. The retaliation could make the envy quite expensive. Instead of worrying about the relative profits of the seller, the buyer should worry about whether another buying strategy would be better.

Like a business partner who never cheats anyone, TIT FOR TAT never beats anyone-yet both do very well for themselves.

One idea that is amazingly counterintuitive at first in the Prisoner's Dilemma is that the best possible strategy to follow is ALL D if the other player is unresponsive. It might seem that some form of random strategy might do better, but that is completely wrong. If I have laid out all my moves in advance, then playing TIT FOR TAT will do you no good, nor will flipping a coin. You should simply defect every move. It matters not what pattern I have chosen. Only if I can be influenced by your play will it ever do you any good to cooperate.

Fortunately, in an environment where there are programs that cooperate (and whose cooperation is based on reciprocity), being unresponsive is a very poor strategy, which in turn means that ALL D is a very poor, strategy. The single unresponsive competitor in the second tournament was RANDOM and it finished next to last. The last-place finishers strategy was responsive, but its behavior was so inscrutable that it looked unresponsive

And in a more recent computer tournament conducted by Marek Lugowski, and myself in the Computer Science Department at Indiana L-university three ALL-D's came in at the very bottom (out of 53), with a couple of RA.VDO.11's giving them a tough fight for the honor.

One way to explain TIT FOR TIT's success is simply to say that it elicits cooperation, via friendly persuasion. Axelrod spells this out as follows:

Part of its success might be that other rules anticipate its presence and are designed to do well with it. Doing well with TIT FOR TAT requires cooperating with it, and this in turn helps TIT FOR TIT Even rules that were designed to see what then could get away with quickly apologize to TIT FOR F-IT. A rule that tries to take advantage of TIT FOR TAT will simply hurt itself. TIT FOR TAT benefits from its own nonexploitability because three conditions are satisfied:

1. The possibility of encountering TIT FOR TAT is salient;
2. Once encountered, TIT FOR TAT is easy to recognize; and
3. Once recognized, TIT FOR TAT's nonexploitability is easy to appreciate

This brings out a fourth "personality trait" (in addition to niceness, provocability, and forgiveness) that may play an important role in success: recognizability, or straightforwardness. Axelrod chooses to call this trait clarity, and argues for it with clarity:

Too much complexity can appear to be total chaos. If you are using a strategy that appears random, then you also appear unresponsive to the other player. If you are unresponsive, then the other player has no incentive to cooperate with you. So being so complex as to be incomprehensible is yet dangerous.

How rife this is with morals for social and political behavior! It is rich food for thought.

* * *

Anatol Rapoport cautions against overstating the advantages of TIT FOR TIT; in particular, he believes that TIT FOR LIT is too harshly retaliatory on occasion. It can also be persuasively argued that TIT FOR TAT is too lenient on other occasions. Certainly there is no evidence that TIT FOR TAT is the ultimate or best possible strategy. Indeed, as has been emphasized repeatedly, the very concept of "best possible" is incoherent, since all depends on environment. In the tournament at Indiana University mentioned earlier, several TIT-FOR-TAT-like strategies did better than pure TIT FOR LIT did. They all shared, however, the three critical "character traits" whose desirability had been so clearly delineated by Axelrod's prior analysis of the important properties of TIT FOR TAT. They were simple a little better than TIT FOR LIT at detecting nonresponsiveness, and when they were convinced the other player was unresponsive, they switched over to an ALL-D mode.

In his book, Axelrod takes pains to spell out the answers to three fundamental questions concerning the temporal evolution of cooperation in a world of raw egoism. The first concerns initial viability: How can cooperation get started in a world of unconditional defection—a "primordial sea" swarming with unresponsive ALL-D creatures? The answer (whose proof I omit here) is that invasion by small clusters of conditionally cooperating organisms, even if they form a tiny minority, is enough to give cooperation a toehold. One cooperator alone will die, but small clusters of cooperators can arrive (via mutation or migration, say) and propagate even in a hostile environment, provided they are defensive like TIT FOR TAT. Complete pacifists—Quaker-like programs—will not survive, however, in this harsh environment.

The second fundamental question concerns robustness: What type of strategy does well in unpredictable and shifting environments? We have already seen that the answer to this question is: Any strategy possessing the four fundamental "personality traits" of niceness, provocability, forgiveness, and clarity. This means that such strategies, once established, will tend to flourish, especially in an ecologically evolving world.

The final question concerns stability: Can cooperation protect itself from invasion? Axelrod proved that it can indeed. In fact, there is a gratifying asymmetry to his findings: Although a world of "meanies" (beings using the inflexible ALL-D strategy) is penetrable by cooperators in clusters, a world of cooperators is not penetrable by meanies, even if they arrive in clusters of any size. Once cooperation has established itself, it is permanent. As Axelrod puts it, "The gear wheels of social evolution have a ratchet."

The term "social" here does not mean that these results necessarily apply only to higher animals that can think. Clearly, four-line computer programs do not think—and yet, it is in a world of just such "organisms" that cooperation has been shown to evolve. The only "cognitive abilities" needed by TIT FOR TAT are: (1) recognition of previous partners, and (2) memory of what happened last time with this partner. Even bacteria can do this, by interacting with only one other organism (so that recognition is automatic) and by responding only to the most recent action of their "partner" (so that memory requirements are minimal). The point is that the entities involved can be on the scale of bacteria, small animals, large animals, or nations. There is no need for "reflective rationality"; indeed, TIT FOR TAT could be called "reflexive" (in the sense of being as simple as a knee-jerk reflex) rather than "reflective".

For people who think that moral behavior toward others can emerge only when there is imposed some totally external and horrendous threat (say, of the fire-and-brimstone sort) or soothing promise of heavenly reward (such as eternal salvation), the results of this research must give pause for thought. In one sentence, Axelrod captures the whole idea: Mutual cooperation can

emerge in a world of egoists without central control, by starting with a cluster of individuals who rely on reciprocity.

There are so many situations in the world today where these ideas seem of extreme relevance—indeed, urgency—that it is very tempting to draw morals all over the place. In the later chapters of his book, Axelrod offers advice about how to promote cooperation in human affairs, and at the very end the political scientist in him cautiously ventures some broad conclusions concerning global issues, which are a fitting way for me to conclude as well.

Today, the most important problems facing humanity are in the arena of international relations where independent, egoistic nations face each other in a state of near anarchy. Many of these problems take the form of an iterated Prisoner's Dilemma. Examples can include arms races, nuclear proliferation, crisis bargaining, and military escalation. Of course, a realistic understanding of these problems would have to take into account many factors not incorporated into the simple Prisoner's Dilemma formulation, such as ideology, bureaucratic politics, commitments, coalitions, mediation, and leadership. Nevertheless, we can use all the insights we can get.

Robert Gilpin [in his book *War and Change in World Politics*] points out that from the ancient Greeks to contemporary scholarship all political theory addresses one fundamental question: "How can the human race, whether for selfish or more cosmopolitan ends, understand and control the seemingly blind forces of history?" In the contemporary world this question has become especially acute because of the development of nuclear weapons.

The advice given in this book to players of the Prisoner's Dilemma might also serve as good advice to national leaders as well: Don't be envious, don't be the first to defect, reciprocate both cooperation and defection, and don't be too clever. Likewise, the techniques discussed in this book for promoting cooperation in the Prisoner's Dilemma might also be useful in promoting cooperation in international politics.

The core of the problem is that trial-and-error learning is slow and painful. The conditions may all be favorable for long-run developments, but we may not have the time to wait for blind processes to move us slowly towards mutually rewarding strategies based upon reciprocity. Perhaps if we understand the process better, we can use our foresight to speed up the evolution of cooperation.

Post Scriptum.

In the course of writing this column and thinking the ideas through, I was forced to confront over and over again the paradox that the Prisoner's Dilemma presents. I found that I simply could not accept the seemingly flawless logical conclusion that says that a rational player in a noniterated situation will always defect. In turning this over in my mind and trying to articulate my objections clearly, I found myself inventing variations after

variation after variation on the basic situation. I would like to describe just a few here.

A version of the dealer-and-buyer scenario involving bags exchanged in a forest actually occurs in a more familiar context. Suppose I take my car in to get the oil changed. I know little about auto mechanics, so when I come in to pick it up. I really have no way to verify if they've done the job. For all I know, it's been sitting untouched in their parking lot all day, and as I drive off they may be snickering behind my back. On the other hand, maybe I've got the last laugh, for how do they know if that check I gave them will bounce?

This is a perfect example of how either of us could defect, but because the situation is iterated, neither of us is likely to do so. On the other hand, suppose I'm on my way across the country and have some radiator trouble near Gillette, Wyoming, and stop in town to get my radiator repaired there. There is a decent chance now that one party or the other will attempt to defect, because this kind of situation is not an iterated one. I'll probably never again need the services of this garage, and they'll never get another check from me. In the most crude sense, then, it's not in my interest to give them a good check, nor is it in theirs to fix my car. But do I really defect? Do I give out bad checks? No. Why not?

Consider this related situation. Late at night, I bang into someone's car in a deserted parking lot. It's apparent to me that nobody witnessed the incident. I have the choice of leaving a note, telling the owner who's to blame, or scurrying off scot-free. Which do I do? Similarly, suppose I have given a lecture in a classroom in a university I am visiting for one day, and have covered the board with chalk. Do I take the trouble of erasing the board so that whoever comes in the next morning won't have to go to that trouble? Or do I just leave it?

I was recently waiting to board an airplane when a voice announced: "Passengers holding seats in rows 24 to 36 may now board." Well, my seat was in row 4, so I waited. A few minutes later, the voice said that passengers in rows 18 to 36 were free to board. A group of people got up and went in. Then after a couple of minutes, rows 10 to 36 were told they could board. A dozen people or so remained in the waiting area. For a while, we were all patient, waiting for the final announcement allowing us to board, but after about five minutes, people started fidgeting a bit and edging up toward the gate. Then, after another two or three minutes, a couple of people just went right on. And then the rest of us wondered, "Should we get on, too? Will we be left behind?" For most of the people, the answer was obvious: they rushed to board. And once they had boarded, then the rest of us felt kind of like suckers, and we just got on too. In effect, there was a stampede that converted cooperators into defectors. Even the people who triggered the stampede had originally been cooperating, but after a while the temptation

got to be too great, and they broke down. At that point, some sort of phase transition, or collective shift, took place, and the stable state of patient cooperation collapsed into a chaotic scrambling for places. Actually, it wasn't that bad, and there was a good reason for the relatively polite way we did board, defectors though we were: we all had seat assignments, so it didn't matter who got on first. But imagine if the earliest defectors were sure to get the best remaining seats! The contemporary aphorist .Ashleigh Brilliant has found just the right *bons mots* to describe this sort of dilemma:

Should I abide by the rules until they're changed, or help speed tht change by breaking them?

Better start rushing before the rush begins!

In pondering the Prisoner's Dilemma, I could not help but be reminded of horrible scenarios in Nazi concentration camps, where large herds of unarmed people would be led to their deaths by small herds of armed people. It seems that a stampede by the masses could quickly have overcome a small number of guards, at least in certain critical narrow passageways here and there. The trouble is, it would require certain death on the part of a few ultra-cooperators. in exchange for the liberation of a large number of other people. Generally speaking, individuals are not willing to perform such an exchange. Nobody wants to be in the front lines of a protest demonstration facing troops with machine guns. Everyone wants to be in the rear. But not everyone can be in the rear! If nobody is willing to be in the front lines, then there will be no front lines, and consequently no demonstration at all.

Driving a car has a certain primitive quality to it that brings out the animal in us all, and probably that's why it confronts us with Prisoner's-Dilemma like situations so often-more often than any other activity I can think of. How about those annoying drivers who, when there's a long line at a freeway exit, zoom by all the politely lined-up cars and then butt in at the very last moment, getting off 50 cars ahead of you? Are you angry t such people. or do you do it too? Or, worse-do you do it and yet resent others who have such gall?

I have been struck by the relative savagery of the driving environment in the Boston area. I know of no other city in which people are so willing to take the law into their own hands, and to create complete anarchy. There seems to be less respect for such things as red lights. stop signs, lines in the street, speed limits, other people's cars, and so forth, than in any other city or state, or country that I have ever driven in. This incessant "me-first" attitude seems to be a vicious, self-reinforcing circle. Since there are so many people who do whatever they want, nobody can afford to be polite and let other

people in ahead of them (say), for then they will be taken advantage of repeatedly and will wind up losing totally. You simply must assert yourself in many situations, and that means you must defect. Of course, just one defection does not an .ALL-D player make. In fact, a retaliatory defection is just good old TIT-FOR-TAT playing. However, very often in Boston driving, there is no way you can get back at a nasty driver who cuts in front of you and then takes off screeching around the corner. That person is gone forever. You can take out your frustrations only on the rest of the people near you, who are not to blame for that driver. You can cut in ahead of them. Does this do any good? That is, does it teach anybody a lesson? Obviously it will teach them only that it pays to defect. And thus the spiral starts.

Is there any way to put a halt to the descending spiral, the vortex towards oblivion? Is there any point at which the people of Boston will collectively come to realize that it has gotten so bad that they will all suddenly "flip" and begin to cooperate in situations where they formerly would have defected? Can there be a stampede toward cooperation, just as there can be a stampede toward defection?

Clearly, if large numbers of people were to start driving much less aggressively and nastily, everybody would benefit. Huge snarls would unsnarl-in fact would never form. Traffic would flow smoothly and regularly. The shoulders-those favorite illegal passing lanes for defectors-would be completely clear. So clear, in fact, that just think-you and I could make sensational progress by swerving onto an empty shoulder and passing everybody. Wheee! Isn't this fun? Aren't those other people suckers, staying in the slow lane and glaring at us? Say, how come other people are barging in on us? This is our lane. Oh, so that person in the yellow car wants to play dirty, eh? Okay. I'll show them what playing dirty's really like!

Sound familiar? Is there any solution to such terrible spirals? Sometimes I am very pessimistic on that subject. Anatol Rapoport and I exchanged letters concerning this matter, and he related a frightening anecdote. I quote from his letter:

Do you know of the experiment performed by Martin Shubik, in which a dollar bill was auctioned off for \$3.40? This was a consequence of a rule (the implications of which dawned on the subjects only when they were already hooked) specifying that while the highest bidder got the dollar, the second-highest bidder would also have to pay what he last bid. It thus became imperative to keep going, since the second-highest bidder (whoever he was at each stage) had progressively more to lose as the bids went up. Are Reagan and Andropov too stupid to see the point?

I believe the "technological imperative" is driving our species to extinction. Ever more horrendous weapons must be produced, simply because it is possible to produce them. Eventually the\, must be used, to justify the insane waste. It thus becomes imperative to seal off the "logic" of the paradigm based on "deterrence", "balance of power", and similar metaphors-to make it in-assailable.

I don't think intelligence plays a part in the vicious cycle of the arms race. The rulers only think they make the decisions. If they were C players, they would not be where they are. If they started to play C while in office, they would be impeached, overthrown, or assassinated. Does this mean that D players are selected for? Possibly in the short run, but not on the time scale of evolution. H. sapiens is apparently not the last word, but for me, a homocentric, this is no consolation.

Pretty- sobering words from one of the leading rational thinkers of our era