

Chapter 3

The Evolution of Strategies in the Iterated Prisoner's Dilemma

3.1 The Problem with Cooperation

Mutual cooperation among groups of organisms is a frequently occurring phenomenon. When such cooperation is of benefit to the cooperating agents, and when the lack of it is harmful to them, the mechanisms by which such cooperation might arise and persist seem straightforward.

There are other types of cooperation, however, that are characterized by the fact that, while cooperating agents do well, any one of them would do better by failing to cooperate. For this sort of case, it is more difficult to explain how group cooperation would arise and persist, for one would expect organisms showing a propensity to cooperate to do less well than their neighbors, leading to a dying out of cooperating tendencies in a large, non-cooperating population.

The Prisoner's Dilemma of game theory is an elegant embodiment of this sort of case. In the Prisoner's Dilemma, two individuals can each either cooperate or defect. The payoff to a player affects its reproductive success. No matter what the other does, the selfish choice of defection yields a higher payoff than cooperation. But if both defect, both do worse than if both had cooperated. Figure 3-1 shows the payoff matrix of the Prisoner's Dilemma used in this study.

In many biological settings, the same two individuals may meet more than once. If an individual can recognize a previous interactant and remember some aspects of the prior outcomes, then the strategic situation becomes an iterated Prisoner's Dilemma. In an iterated Prisoner's Dilemma, a strategy is a decision rule which specifies the probability of cooperation or defection as a function of the history of the interaction so far.

To see what type of strategy can thrive in a variegated environment of more or less sophisticated strategies, I conducted a computer tournament for the iterated Prisoner's Dilemma. The strategies were submitted by game theorists in economics, sociology, political science, and mathematics (Axelrod, 1980a). The 14 entries and a totally random strategy were paired with each other in a round robin tournament. Some of the strategies were quite intricate. An example is one which on each move models the behavior of the other player as a Markov process, and then uses Bayesian inference to select what seems the best choice for the long run. However, the result of the tournament

		Column Player	
		Cooperate	Defect
Row Player	Cooperate	$R = 3, R = 3$ Reward for mutual cooperation	$S = 0, T = 5$ Sucker's payoff, and temptation to defect
	Defect	$T = 5, S = 0$ Temptation to defect and sucker's payoff	$P = 1, P = 1$ Punishment for mutual defection

Note: The payoffs to the row chooser are listed first.

Figure 3-1: The Prisoner's Dilemma

was that the highest average score was attained by the simplest of all strategies, TIT FOR TAT. This strategy is simply one of cooperating on the first move and then doing whatever the other player did on the preceding move. TIT FOR TAT is a strategy of cooperation based upon reciprocity.

The results of the first round were circulated and entries for a second round were solicited. This time there were 62 entries from six countries (Axelrod, 1980b). Most of the contestants were computer hobbyists, but there were also professors of evolutionary biology, physics, and computer science, as well as the five disciplines represented in the first round. TIT FOR TAT was again submitted by the winner of the first round, Anatol Rapoport. It won again.

The second round of the computer tournament provides a rich environment in which to test the evolution of behavior. It turns out that just eight of the entries can be used to account for how well a given rule did with the entire set. These eight rules can be thought of as representatives of the full set in the sense that the scores a given rule gets with them can be used to predict the average score the rule gets over the full set. In fact, 98% of the variance in the tournament scores is explained by knowing a rule's performance with these eight representatives. So these representative strategies can be used as a complex environment in which to evaluate an evolutionary simulation. That is the environment I used to simulate the evolution of strategies for the iterated Prisoner's Dilemma.

3.2 The Genetic Algorithm

The inspiration for my simulation technique came from an artificial intelligence procedure developed by computer scientist John Holland (Holland 1975, 1980, 1986). Holland's technique is called the genetic algorithm. Using a genetic algorithm, one represents strategies as chromosomes. Each chromosome serves a dual purpose: it provides a representation of what the organism will become, and it also provides the actual material which can be transformed to yield new genetic material for the next generation.

Before going into details, it may help to give a brief overview of how the genetic algorithm works. The first step is to specify a way of representing each allowable strategy as a string of genes on a chromosome which can undergo genetic transformations, such as mutation. Then the initial population is constructed from the allowable set (perhaps by simply picking at random). In each generation, the effectiveness of each individual in the population is determined by running the individual in the current strategic environment. Finally, the relatively successful strategies are used to produce offspring which resemble the parents. Pairs of successful offspring are selected to mate and produce the offspring for the next generation. Each offspring draws part of its genetic material from one parent and part from another. Moreover, completely new material is occasionally introduced through mutation. After many generations of selection for relatively successful strategies, the result might well be a population that is substantially more successful in the given strategic environment than the original population.

To explain how this works, consider the strategies available for playing the iterated Prisoner's Dilemma. In particular, consider the set of strategies that are deterministic and use the outcomes of the three previous moves to make a choice in the current move. Since there are four possible outcomes for each move, there are $4 \times 4 \times 4 = 64$ different histories of the three previous moves. Therefore to determine its choice of cooperation or defection, a strategy would only need to determine what to do in each of the situations which could arise. This could be specified by a list of sixty-four C's and D's (C for cooperation and D for defection). For example, one of these sixty-four genes indicates whether the individual cooperates or defects when in a rut of three mutual defections. Other parts of the chromosome would cover all the other situations that could arise.

To get the strategy started at the beginning of the game, it is also necessary to specify its initial premises about the three hypothetical moves which preceded the start of the game. To do this requires six more genes, making a total of seventy loci on the chromosome.¹ This string of seventy C's and D's would specify what the individual would do in every possible circumstance and would therefore completely define a particular strategy. The string of 70 genes would also serve as the individual's chromosome for use in reproduction and mutation.

There is a huge number of strategies which can be represented in this way. In fact, the number is 2^{70} , which is about 10^{21} .² An exhaustive search for good strategies in this huge collection of strategies is clearly out of the question. If a computer had examined these strategies at the rate of 100 per second since the beginning of the universe, less than one percent would have been checked by now.

To find effective strategies in such a huge set, a very powerful technique is needed.

¹The six premise genes encode the presumed C or D choices made by the individual and the other player in each of the three moves before the interaction actually begins.

²Some of these chromosomes give rise to equivalent strategies since certain genes might code for histories that could not arise given how loci are set. This does not necessarily make the search process any easier, however.

Holland's "genetic algorithm" is such a technique. Genetic algorithms were originally inspired by biological genetics, but were adapted by Holland to be a general problem-solving technique. In the present context, a genetic algorithm can be regarded as a model of a "minimal genetics" which can be used to explore theoretical aspects of evolution in rich environments. The simulation program works in five stages:

1. An initial population is chosen. In the present context the initial individuals can be represented by random strings of seventy C's and D's.
2. Each individual is run in the current environment to determine its effectiveness. In the present context this means that each individual player uses the strategy defined by its chromosome to play an iterated Prisoner's Dilemma with other strategies, and the individual's score is its average over all the games it plays.³
3. The relatively successful individuals are selected to have more offspring. The method used is to give an average individual one mating, and to give two matings to an individual who is one standard deviation more effective than the average. An individual who is one standard deviation below the population average would then get no matings.
4. The successful individuals are then randomly paired off to produce two offspring per mating. For convenience, a constant population size is maintained. The strategy of an offspring is determined from the strategies of the two parents. This is done by using two genetic operators: crossover and mutation.
 - a. Crossover is a way of constructing the chromosomes of the two offspring from the chromosomes of two parents. It can be illustrated by an example of two parents, one of whom has seventy C's in its chromosome (indicating that it will cooperate in each possible situation that can arise), and the other of whom has seventy D's in its chromosome (indicating that it will always defect). Crossover selects one or more places to break the parents' chromosomes in order to construct two offspring each of whom has some genetic material from both parents. For example, if a single break occurs after the third gene, then one offspring will have three C's followed by sixty-seven D's, while the other offspring will have three D's followed by sixty-seven C's.
 - b. Mutation in the offspring occurs by randomly changing a very small proportion of the C's to D's or vice versa.
5. This gives a new population. This new population will display patterns of behavior that are more like those of the successful individuals of the previous generation, and

³The score is actually a weighted average of its scores with the eight representative, the weights having been chosen to give the best representation of the entire set of strategies in the second round of the tournament.

less like those of the unsuccessful ones. With each new generation, the individuals with relatively high scores will be more likely to pass on parts of their strategies, while the relatively unsuccessful individuals will be less likely to have any parts of their strategies passed on.

3.3 Simulation Results

The computer simulations were done using a population size of twenty individuals per generation. Levels of crossover and mutation were chosen averaging one crossover and one-half mutation per chromosome per generation. Each game consisted of 151 moves, the average game length used in the tournament. With each of the twenty individuals meeting eight representatives, this made about 24,000 moves per generation. A run consisted of 50 generations. Forty runs were conducted under identical conditions to allow an assessment of the variability of the results.

The results are quite remarkable: from a strictly random start, the genetic algorithm evolved populations whose median member was just as successful as the best rule in the tournament, TIT FOR TAT. Most of the strategies that evolved in the simulation actually resemble TIT FOR TAT, having many of the properties that make TIT FOR TAT so successful. For example, five behavioral alleles in the chromosomes evolved in the vast majority of the individuals to give them behavioral patterns that were adaptive in this environment and mirrored what TIT FOR TAT would do in similar circumstances. These patterns are:

1. Don't rock the boat: continue to cooperate after three mutual cooperations (which can be abbreviated as C after RRR).
2. Be provokable: defect when the other player defects out of the blue (D after receiving RRS).
3. Accept an apology: continue to cooperate after cooperation has been restored (C after TSR).
4. Forget: cooperate when mutual cooperation has been restored after an exploitation (C after SRR).
5. Accept a rut: defect after three mutual defections (D after PPP).

The evolved rules behave with specific representatives in much the same way as TIT FOR TAT does. They did about as well as TIT FOR TAT did with each of the eight representatives. Just as TIT FOR TAT did, most of the evolved rules did well by achieving almost complete mutual cooperation with seven of the eight representatives. Like TIT FOR TAT, most of the evolved rules do poorly with only one representative.

called ADJUSTER, that adjusts its rate of defection to try to exploit the other player. In all, 95% of the time the evolved rules make the same choice as TIT FOR TAT would make in the same situation.

While most of the runs evolve populations whose rules are very similar to TIT FOR TAT, in eleven of the forty runs, the median rule actually does substantially better than TIT FOR TAT.⁴ In these eleven runs, the populations evolved strategies that manage to exploit one of the eight representatives at the cost of achieving somewhat less cooperation with two others. But the net effect is a gain in effectiveness.

This is a remarkable achievement because to be able to get this added effectiveness, a rule must be able to do three things. First, it must be able to discriminate between one representative and another based upon only the behavior the other player shows spontaneously or is provoked into showing. Second, it must be able to adjust its own behavior to exploit a representative that is identified as an exploitable player. Third, and perhaps most difficult, it must be able to achieve this discrimination and exploitation without getting into too much trouble with the other representatives. This is something that none of the rules originally submitted to the tournament were able to do.

These very effective rules evolved by breaking the most important advice developed in the computer tournament, namely to be "nice", that is never to be the first to defect. These highly effective rules always defect on the very first move, and sometimes on the second move as well, and use the choices of the other player to discriminate what should be done next. The highly effective rules then had responses that allowed them to "apologize" and get to mutual cooperation with most of the unexploitable representatives, and they had different responses which allowed them to exploit a representative that was exploitable.

While these rules are highly effective, it would not be accurate to say that they are better than TIT FOR TAT. While they are better in the particular environment consisting of fixed proportions of the eight representatives of the second round of the computer tournament, they are probably not very robust in other environments. Moreover, in an ecological simulation these rules would be destroying the basis of their own success, as the exploited representative would become a smaller and smaller part of the environment (Axelrod 1984, pp. 49-52 and 203-5). While the genetic algorithm was sometimes able to evolve rules that are more effective than any entry in the tournament, the algorithm was only able to do so by trying many individuals in many generations against a fixed environment. In sum, the genetic algorithm is very good at what actual evolution does so well: developing highly specialized adaptations to specific environmental settings.

In the evolution of these highly effective strategies, the computer simulation employed sexual reproduction, where two parents contributed genetic material to each offspring. To see what would happen with asexual reproduction, forty additional runs were conducted in which only one parent contributed genetic material to each offspring. In

⁴The criterion for being substantially better than TIT FOR TAT is a median score of 450 points, which compares to TIT FOR TAT's weighted score of 428 with these eight representatives.

these runs, the populations still evolved toward rules that did about as well as TIT FOR TAT in most cases. However, the asexual runs were only half as likely to evolve populations in which the median member was substantially more effective than TIT FOR TAT.⁵

So far, the simulation experiments have dealt with populations evolving in the context of a constant environment. What would happen if the environment is also changing? To examine this situation, another simulation experiment with sexual reproduction was conducted in which the environment consisted of the evolving population itself. In this experiment each individual plays the iterated Prisoner's Dilemma with each other member of the population rather than with the eight representatives. At any given time, the environment can be quite complex. For an individual to do well requires that its strategy achieves a high average effectiveness with the nineteen other strategies that are also present in the population. Thus as the more effective rules have more offspring, the environment itself changes. In this case, adaptation must be done in the face of a moving target. Moreover, the selection process is frequency dependent, meaning that the effectiveness of a strategy depends upon what strategies are being used by the other members of the population.

The results of the ten runs conducted in this manner display a very interesting pattern. From a random start, the population evolves away from whatever cooperation was initially displayed. The less cooperative rules do better than the more cooperative rules because at first there are few other players who are responsive — and when the other player is unresponsive the most effective thing for an individual to do is simply defect. This decreased cooperation in turn causes everyone to get lower scores as mutual defection becomes more and more common. However, after about ten or twenty generations the trend starts to reverse. Some players evolve a pattern of reciprocating what cooperation they find, and these reciprocating players tend to do well because they can do very well with others who reciprocate without being exploited for very long by those who just defect. The average scores of the population then start to increase as cooperation based upon reciprocity becomes better and better established. So the evolving social environment led to a pattern of decreased cooperation and decreased effectiveness, followed by a complete reversal based upon an evolved ability to discriminate between those who will reciprocate cooperation and those who won't. As the reciprocators do well, they spread in the population resulting in more and more cooperation and greater and greater effectiveness.

3.4 Lessons

1. The genetic algorithm is a highly effective method of problem solving. Following

⁵This happened in 5 of the 40 runs with asexual reproduction compared to 11 of the 40 runs with sexual reproduction. This difference is significant at the .05 level using the one tailed chi-squared test.

Quincy Wright (1977, pp. 452-454), the problem for evolution can be conceptualized as a search for relatively high points in a multidimensional field of gene combinations, where height corresponds to fitness. When the field has many local optima, the search becomes quite difficult. When the number of dimensions in the field becomes great, the search is even more difficult. What the computer simulations demonstrate is that the genetic algorithm is a highly efficient method for searching such a complex multidimensional space. The first experiment shows that even with a seventy dimensional field of genes, quite effective strategies can be found within fifty generations. Sometimes the genetic algorithm found combinations of genes that violate the previously accepted mode of operation (not being the first to defect) to achieve even greater effectiveness than had been thought possible.

2. Sexual reproduction does indeed help the search process. This was demonstrated by the much increased chance of achieving highly effective populations in the sexual experiment compared to the asexual experiment. If sexual reproduction comes at the cost of reduced fecundity, it is not clear whether this gain in search efficiency would be worth the cost of fewer offspring. However, one case in which search efficiency can be very important is the escape from rapidly evolving parasites. This has been demonstrated in one and two locus models (Hamilton, 1980) and shown to be relevant in the sexual selection of birds (Hamilton, 1982).
3. Some aspects of evolution are arbitrary. In natural settings, one might observe that a population has little variability in a specific gene. In other words one of the alleles for that gene has become fixed throughout the population. One might be tempted to assume from this that the allele is more adaptive than any alternative allele. However, this may not be the case. The simulation of evolution allows an exploration of this possibility by allowing repetitions of the same conditions to see just how much variability there is in the outcomes. In fact, the simulations show two reasons why convergence in a population may actually be arbitrary.
 - a. Genes that do not have much effect on the fitness of the individual may become fixed in a population because they "hitch-hike" on other genes that do (Maynard Smith and Haigh, 1974). For example, in the simulations some sequences of three moves may very rarely occur, so what the corresponding genes dictate in these situations may not matter very much. However, if the entire population are descendants of just a few individuals, then these irrelevant genes may be fixed to the values that their ancestors happened to share. Repeated runs of a simulation allow one to notice that some genes become fixed in one population but not another, or that they become fixed in different ways in different populations.
 - b. In some cases, some parts of the chromosome are arbitrary in content, but what is not arbitrary is that they be held constant. By being fixed, other

parts of the chromosome can adapt to them. For example, the simulations of the individual chromosomes had six genes devoted to coding for the premises about the three moves that preceded the first move in the game. When the environment was the eight representatives, the populations in different runs of the simulation developed different premises. Within each run, however, the populations were usually very consistent about the premises: the six premise genes had become fixed. Moreover, within each population these genes usually became fixed quite early. It is interesting that different populations evolved quite different premises. What was important for the evolutionary process was to fix the premise about which history is assumed at the start so that the other parts of the chromosome could adapt on the basis of a given premise.

4. There is a tradeoff in evolution between the gains to be made from flexibility and the gains to be made from commitment and specialization. Flexibility might help in the long run, but in an evolutionary system, the individuals also have to survive in the short run if they are to reproduce. This feature of evolution arises at several levels.
 - a. As the simulations have shown, the premises became fixed quite early. This meant a commitment to which parts of the chromosome would be consulted in the first few moves, and this in turn meant giving up flexibility as more and more of the chromosome evolved on the basis of what had been fixed. This in turn meant that it would be difficult for a population to switch to a different premise. So flexibility was given up so that the advantages of commitment could be reaped.
 - b. There is also a tradeoff between short and long term gains in the way selection was done in the simulation experiments. In any given generation there would typically be some individuals that did much better than the average, and some that did only a little better than the average. In the short run, the way to maximize the expected performance of the next generation would be to have virtually all of the offspring come from the very best individuals in the present generation. But this would imply a rapid reduction in the genetic variability of the population, and a consequent slowing of the evolutionary process later on. If the moderately successful were also given a chance to have some offspring, this would help the long term prospects of the population at the cost of optimizing in the short run. Thus there is an inherent tradeoff between exploitation and exploration, i.e. between exploiting what already works best and exploring possibilities that might eventually evolve into something even better (Holland, 1975, p. 160).

3.5 Conclusions

The genetic simulations provided in this paper are highly abstract systems. The populations are very small, and the number of generations is few. More significantly, the genetic process have only two operators, mutation and crossover, and the sexual reproduction has no sexual differentiation and always had two offspring per mating. These are all highly simplified assumptions, and yet the simulations displayed a remarkable ability to evolve sophisticated adaptive strategies in moderately complex environments.

In the future, more complex and realistic simulations are possible. But the main advantage of simulations can already be glimpsed from these minimal simulation experiments. They provide a different intellectual perspective on evolution. Instead of having to rely only on our observations of real biological systems or our standard mathematical models, we will be able to approach genetics and evolution as a theoretical design problem. We can begin asking about whether parasites are inherent in all complex systems, or are merely the outcome of the way biological systems have happened to evolve. We can begin investigating alternative ways genetics might have evolved and see just which properties of our biological heritage are arbitrary and which are not. Today microbiologists are developing the techniques to alter our genetic heritage. Perhaps now is also the time to think about doing some "as if" experiments to better appreciate the fundamental properties of the genetic system that is the basis of our natural endowment.

3.6 Acknowledgements

I thank Stephanie Forrest and Reiko Tanese for their help with the computer programming, Michael D. Cohen and John Holland for their helpful suggestions, and the Harry Frank Guggenheim Foundation and the National Science Foundation for their financial support to Robert Axelrod.