Chapter 12

Altruism

One of the hardest features of animal behavior to explain is altruism because it seems to fly in the face of fitness-driven evolution. Simple forms of altruism can be explained by the usefulness of protecting ones genetic investment in relatives other than direct offspring. For more complicated models we have to turn to game theory which shows that altruistic behavior can be optimal but the conditions under which it is so are delicate. In any case we do better when we can estimate average behavior over trials. One shot decisions are risky and problematic.

12.1 Protecting your genes

In the last chapter you saw that the computation done by genetic algorithms provides a basis for thinking about hominids. It turns out that the viewpoint of the gene helps us to think about altruism as well. This view was the brainchild of William Hamilton. His idea was that individuals would invest in relatives to the extent of the amount of genes that they shared with that relative. Identical twins have identical DNA, but a child has $\frac{1}{2}$ of the DNA of the parent, and the grandchild has $\frac{1}{4}$.

It is hard to imagine how confusing things were before Hamilton's clarifying work. Now things are not only much clearer, but we can appreciate a revolution in thinking. As Dawkins has eloquently portrayed [2], in the long view we should think of the genes as in charge, using the phenotype as a way to get copies of themselves into the next generation.

But while the genetic view explains a lot of data it does not explain

all of altruism. While in I was in India a girl of twelve and her younger brother had escaped the path of a speeding car when the girl looked back to see her neighbor's child still in danger. She ran back and pushed him out of the way but was hit by the car and lost her leg in the process. Since she and her neighbor did not share any genes we have to look elsewhere for explanations of such cases. It turns out that we can begin to model these cases with computation and the specialty that does so is called *Game Theory*. Such models do not settle all the complexities of social interactions by a long shot. But they do introduce the important idea that cooperation is a rational thing to do. And if that its true then it is entirely possible that the need to cooperate is specified in the genes. It is extremely likely that we are hard-wired to be good.

12.2 Social intelligence and Game Theory

At the height of the Cold War between the US and USSR, the specter of a nuclear holocaust prompted the best thinkers to try and come up some sort of escape plan. While a solution was never found, Morgenstern and von Neumann turned to game theory for a way of thinking about the problem. In game theory, the negotiations between two sides are presented in terms of a table of the various combinations of choices open to both sides. There can be an arbitrary number of sides, but in all our examples we will only use two with players Alice and Bob. In our first example, Alice and Bob each have options C and D. Table 12.1 represents the value to each of the players for different combinations of choices. The way the game is played is that both players make their choices without knowing in advance what the other chose. The choices are then revealed and the payoff matrix consulted to award points to each player.

In this example, Alice should choose option C and Bob should also option C as that results in the highest payoff for both players. This is also a *Nash equilibrium* point as neither Alice or Bob can improve on these choices assuming the other player does what is best for him or her.

The model for the Cold War stalemate is more like our second example: The famous Prisoner's Dilemna game. In this game, summarized in Table 12.2 two prisoners each have the chance of receiving significant rewards for cooperating with the authorities, provided their confederate refuses. If they both rat on each other, the rewards are minimal and if they both refuse

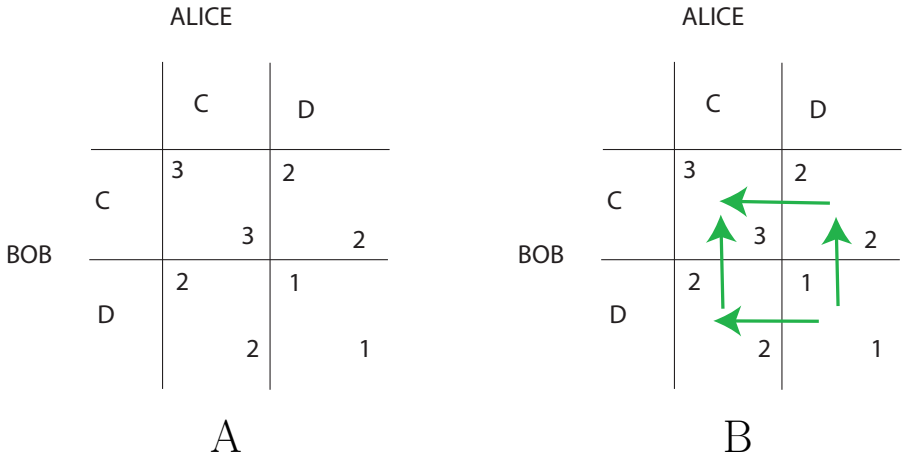


Figure 12.1: A) A very simple game in which if Alice and Bob both choose 'C,' they each receive 3 points etc., B) If they know the payoffs, they can reason that they should each choose this option.

to turn in their confederate, the rewards are modest. The payoff matrix looks like this:

In the slightly confusing jargon of this game, not turning in your confederate is termed "cooperating" and ratting on him or her is termed "defecting." What should Alice do? If Bob is going to defect, then from the table, the best Alice can do is to defect as well. She'll at least get 1 point. If Bob is going to cooperate, then by defecting she'll get 5 points compared to only 3 for cooperating as well. So no matter what Bob does the rational thing to do is defect. Naturally the same logic holds for Bob as well. Thus it seems that the players ineevitably much each choose 'defect.' Interpreting this example in terms of the Cold War, it seemed as though the right thing for the US or USSR to do was to "defect" and shoot their missiles first. A lot of people on both sides advocated doing just this at the time. For the West, one infamous example was most Bertrand Russell[4].

Although this would not work for the Cold War situation, the situation for Alice and Bob changes if they are recidivists. They play the game, see the choice that each has made and then play again. This version of the game is known as the *Iterated Prisoner's Dilemna* or IPD. If they are going to play the game over and over again, there is the possibility that they could find the compromise strategy where they both cooperate. To see that there

		Alice C	Alice D
	C	3 / 3	5 / 0
Bob	D	0 / 5	1 / 1

Table 12.1: The Prisoner's Dilemma Game. Alice and Bob are hardened criminals who are caught and separately asked to turn their partner in return for years taken off their sentence. If they cooperate (C) that is refuse to tell then the sentence possible is three years below maximum. Alice's reward is shown in the upper left of each box, Bob's reward is shown in the lower right. So if Bob 'defects' i.e turns in Alice, while Alice cooperates, he receives a five year sentence reduction and she none.

might be good prospects, we can analyze the outcomes where a defector and a cooperator each play against a player using the strategy "tit-for-tat," abbreviated TFT. The tit-for-tat strategist cooperates for one move and then chooses the opponent's most recent choice move for the subsequent turns. Owing to the uncertainty in the opponent's choices, let's use a discount factor γ that devalues future rewards.

The cooperator will get

$$3 + 3\gamma + 3\gamma^2 + 3\gamma^3 + \cdots = \frac{3}{1-\gamma}$$

And the defector will get

$$5 + \gamma + \gamma^2 + \gamma^3 + \cdots = 5 + \frac{\gamma}{1-\gamma}$$

By comparing these two values, you can show that when γ is greater than $\frac{1}{2}$ the cooperator does better. Thus if you value future rewards, its rational to cooperate, but if you do not you should defect.

The pioneering studies introducing IPD were done by Axelrod[1]. To see if cooperative strategies would appear he had groups of human players play against each other. Each player was free to choose his or her own strategy. For example you might tolerate two defections in a row, but then you defect also to send a message. Or you might cooperate most of the time and defect

once in a while to get a profit. Each player was free to choose his or her own strategy against any player at any time. Surprisingly, it turned out that the straightforward tit-for tat strategy beat all rivals.

However the initial tests with human subjects did not allow for too much subtlety and left the thought that perhaps it might be possible to do better with a computational approach that was capable of remembering lots of information about the strategies of individual players. To test this idea, Axelrod turned to a genetic algorithm. The genetic algorithm was trained using a three-move history of the encounters with individuals. An element of the specific encoding can be visualized as shown in Figure 12.2A.

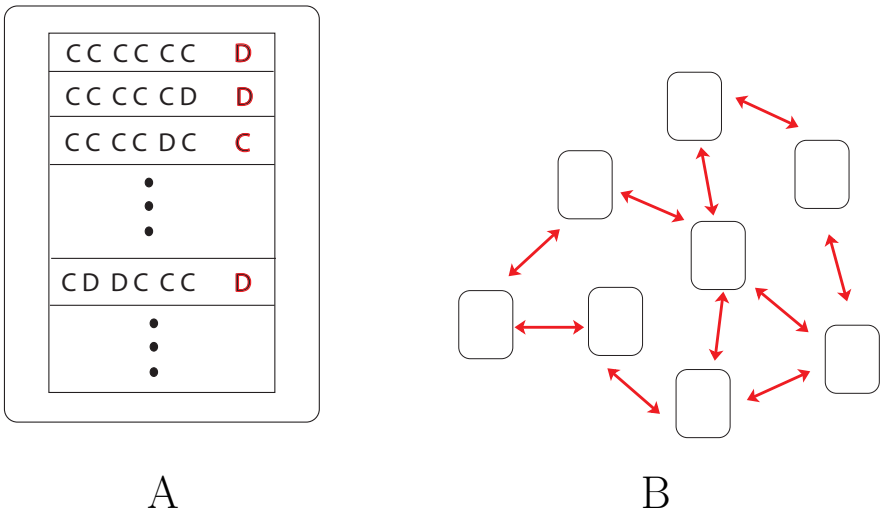


Figure 12.2: A) The particular coding of one of the individuals in the population. The black letters on a row represent possible moves for both players in the most recent three turns. The red letters represent the policy for that individual. B) The genetic algorithm finds good policies by having individuals play against each other to compute a fitness score for each individual.

To understand the encoding. consider the third line of Figure 12.2A. This is interpreted as: if three turns ago I and my opponent both chose cooperate (C), and if two turns ago we also cooperated, and in the penultimate turn my opponent chose D while I cooperated, this time I should choose C. With all possible combinations of the last three turns the 'DNA' string gets quite long since the different strategies need to be represented by different values

in each of the 2^6 possible red positions. As a consequence of this encoding strategy, there is a vast number of *possible* individuals. The total number of such summary strategies, where each has a unique string is 2^{2^6} a large number indeed.

The mechanics of tuning the genetic algorithm were as described in the last chapter. To pick the GA player, a population of individuals each with different such strategies compete against strategies used by human players and evolve according to how well the individuals perform.

Axelrod showed that the GAs beat the TFT strategy. Since TFT was already bettered its human opponents, you might wonder how the GA did it. The secret is that while TFT is a best strategy against optimal opponents, more run-of-the-mill opponent strategies typically have weaknesses that can be exploited in ways that TFT does not take advantage of. If you know your opponent allows two defects before punishing you, you can sneak just one defect in for a profit before behaving.

12.3 IPD and genetics - snail darters

When we see humans playing games that are advertised as being a model of their deep-seated altruistic behavior, computation or not, we tend to be skeptics. We are great at mastering abstractions and playing according to made-up rules, so perhaps this is something we can do but may not do as part of ingrained behavior. To show this view is short-sighted, we turn to the animals that play forms of IPD.

A snail darter is a small fish that schools. When a large fish appears nearby there is always the possibility that that fish is a predator. But also there is the possibility that it is not. The school wants to react appropriately to a predator but not to waste energy on a false alarm. The solution adopted is to allow a pair of scouts approach the predator. As they get close if it is dangerous the predator should reveal itself. But which of the pair should go first? The solution adopted is a form of IPD. Each fish advances by a discrete small amount in turn; first one taking the lead and then the other. The turns are of course like turns in IPD. Choosing to cooperate means going ahead. Turning tail is defecting. By behaving in this way, the two fish share the risk of information gathering.

Tit-for-tat is a good strategy for playing IPD but all versions of it can get jammed if opponents pick to many ‘defects,’ or pick ‘defect’ at the wrong

		Boss: Inspect	Boss: Don't Inspect
	Work	P-W-I W-E	P-W W-E
Worker	Shirk	-I 0	-W W

Table 12.2: The Work or Shirk Game. If the Boss Inspects and the Workers are at it the Boss receives the value of the product (P) minus the wages for the work(W) and the cost of inspecting (I).The worker gets the wages W minus the energy expended (E).

time. For example suppose two TFT strategies play against each other where one TFT strategy is coded as; 'Do whatever your opponent did last time.' If they both cooperate, that is great, but if the opponent ever picks 'defect' then its easy to get stuck in defect forever. The secret to breaking out of this rut is to have probabilistic strategies. You do not have a rigid strategy, but on each round pick your move with a given probability.

12.4 Monkeys do it too

If snail darters are genetically programmed to exhibit cooperative behaviors, then obviously primates should be too. But the question is do they? And if they do, what parts of the brain are implementing these algorithms? Recent work has been done at several laboratories in order to answer these questions.

Glimcher has done experiments using a game called Work or Shirk[3]. In this case the computer plays the role of overseer of a factory that pays a worker a wage W. If the worker, played by the monkey, shows up, a product worth P is produced. The worker would like to skip but if he does, he will loose his wage. If the boss inspects, it costs her I, and if the worker, shows up it cost him E. These data allow us to assemble the payoff matrix describing the different payoffs, and this is shown in Table 12.2

Naturally the worker would like to not show up and be paid while the boss would like to not have to pay the cost of inspection but if either of them do this they can be penalized by the other. The solution is to adopt a probabilistic strategy. It turns out the the best values to each side are

respectively: the probability of inspecting should be $\frac{E}{W}$ and the probability of shirking should be $\frac{I}{W}$. Glimcher showed that monkeys can not only play this game successfully, but that they are sensitive to shifting the payoff values. So if the ratio of I to W is changed then the monkey adjusts his probability accordingly. If the cost of inspecting goes up, then there should be more shirking and this is what happens.

Another simple games is Matching Pennies. Two players turn over a coin at the same time. They can choose the side to have face up. If the sides match then Player A gets a point and Player B gets minus a point. If they do not the payoffs are reversed. This very simple game has some subtleties however. If Player A chooses more heads than tails, and Player B notices this then B can win points by choosing more tails on average. To counter this A should randomly choose heads or tails. Then there is no strategy that can win. However if A does this reliably, then any strategy that B picks is as good as any other. For example if she picks heads all the time the average payoff is still zero.

Lee has trained monkeys to play a version of matching pennies. Using real pennies with monkeys would be a trial since a lot of time would be lost in training the monkeys not to throw the coins or try and eat them. The solution is to have the monkey use eye movements to indicate choices. What happens is that the monkey stares at a light spot in the center of a display and then when the light goes out, looks to one of two lights displayed a short distance either side. The choice of left or right is equivalent to the heads/tails coin decision. The monkey is doing this for a juice reward. Before he looks, the computer has already selected left or right as the reward location. So if the monkey matches this by looking then he actually gets the payoff, otherwise not and a new trial begins.

What is most interesting about this experiment is its control case. In this case the computer selects reward locations from trial to trial at random. So as we just discussed, all strategies that the monkey can try are equivalent. This shows up in the data of the monkey's responses which are arbitrarily biased to a side he prefers. But in the test trials, the computer keeps track of the monkey's responses and tries to minimize his payoff. So if he chooses a lot of rights, the computer choose left more often and vice versa. In this case the monkey must be random in his choices to maximize rewards and that is exactly what he does.

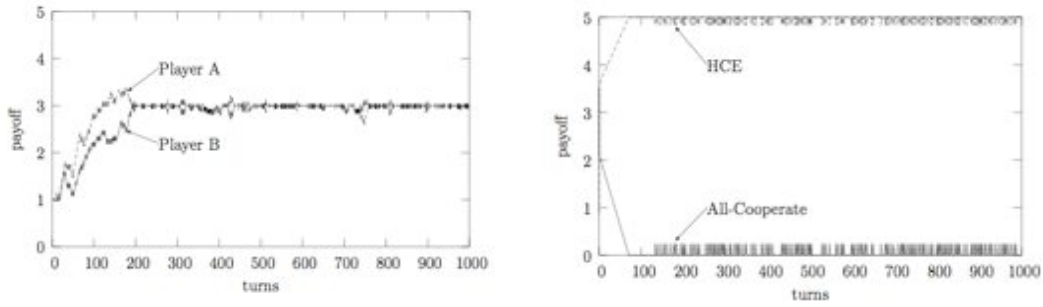
So we have gone way beyond snail darters haven't we? The new factor exhibited by the monkeys is their use of probability in making their choices.

To review, as we saw in the discussion of the standard neural model, when there are unknown payoffs with uncertainty in their estimates, the best thing to do is to sample the one you think is best increasingly often and in the limit you'll have a deterministic strategy. This is the classical Bandit Problem result. Its when these payoffs are changing that it becomes trickier. If they do you have to keep sampling what you estimate are the loosing options more frequently. So you have a probabilistic strategy where your probabilities reflect the continuing uncertainty in the environment. Game Theory can be seen as at the most volatile end of the uncertainty strategy spectrum. Not only is the payoffs varying, but the reason they are varying is that they are being guided by another agent that is trying to exploit you! In these cases, the best defense is a purely probabilistic strategy that is tuned to your opponent's choices. The fact that monkeys play these games optimally is incredibly revealing. Not only does it say that they can do this, and are genetically programmed with these abilities, but that they are tuned to the social demands of dealing with other monkeys in a computationally optimal way.

While the monkeys can handle Work or Shirk, IPD still presents a formidable obstacle as learning to cooperate is a much more delicate problem for the simple reason that the desired result is not a Nash equilibrium. However recently Zhu has shown a way of getting to the desired mutual cooperation point. What he was able to show is that for the case where players are choosing their moves probabilistically, it is possible to test the observations to see whether an opponent is testing the waters by upping the frequency of cooperates. We won't go into the math which is involved, but a player can test to see if the change in the fraction of cooperates is increasing. If this test reports in the affirmative, then the player tries to up his or her fraction but with a certain probability. Think of playing a game of reporting 'Good' or 'Bad,' based on these fractional measurements. It turns out that the game based on fractions is better behaved and can steer a cooperator to the desired point as shown in Figure 12.3A or exploit a die-hard cooperator as in Figure 12.3B

12.5 Populations of Game Players

Finally lets change the venue and look at populations of individuals. This venue has special advantages in that we can explore the dynamics of group



A —LargeB

Figure 12.3: Zhu's model, called HCE for detecting potential cooperators in the IPD works both for formidable opponents A and patsies B.

		Alice Rock	Paper	Scissors
	Rock	0 0	1 -1	-1 1
Bob	Paper	-1 1	0 0	1 -1
	Scissors	1 -1	-1 1	0 0

Table 12.3: The payoffs for the Rock-Paper-Scissors game.

interactions as well as model the spatial locality of groups. To show the properties of probabilistic strategies in populations let's use the setting of the familiar Rock-Paper-Scissors game. As any school kid knows, both players start with their hand behind the back and on command produce a hand configuration that is one of *Rock* (a fist), *Paper* (hand open) or *Scissors* (index and middle finger extended). Rock beats Scissors, Paper beats Rock and Scissors beats Paper. We can represent this as a game with payoffs of one for winning and minus one for losing, as shown in Table12.3.

Now think of a population of Rock-Paper-Scissors players. We can model their behavior in terms of the fraction of the total number of players that are playing any particular strategy. What happens is that at any round, they randomly choose someone to play with and then use their strategy.

Lets assume that the players operate as follows: if any particular strategy is preferred, other players will gradually change their choices to countermand that strategy. So if lots of players are picking 'Rock,' other players will switch to 'Scissors,' and so on. How do you model the dynamics of this behavior? One way do this is to write an equation that expresses a rational for their to be increases in a particular faction. For example, let's model the fraction of players choosing 'Rock' with the equation

$$\dot{r} = r(s - p)$$

In this equation the rate of change of players choosing 'Rock' is denoted by $\dot{r}$ (It could have just as easily been written as $\frac{dr}{dt}$, but the 'dot' notation is nice and compact). What the equation is saying is that if the Scissors players outnumber the Paper players, the choosing Rock is a good thing to do and the proportion of Rock players will increase. Furthermore the rate of increase is proportional to the difference between the fraction of scissors players and the fraction of rock players. Similarly for the other players there is

$$\dot{s} = s(p - r)$$

and

$$\dot{p} = p(r - s)$$

So the model predicts that the proportions will change unless the rates are all zero, i.e.,

$$\dot{r} = \dot{p} = \dot{s} = 0$$

For which values of r, p and s does this happen? You can see right away from these equations that one possibility is

$$r = s = p$$

and since these fractions have to add up to unity, this occurs when they are all equal i.e.,

$$r = s = p = \frac{1}{3}.$$

Because once the fractions achieve these values they will not change, this set of fractions is termed an *equilibrium point*. If the fractions get to this point they will stay there. Of course this is an idealization of the real world as one cannot expect that the fractions will remain perfectly equal. Thus it is useful

to analyze what would happen if they are somehow moved off of this point by a little bit. The possibilities are that the equilibrium point would be unstable or stable. In the former case, once the fractions drift off of the equilibrium point by even a little bit they keep going. In the latter they oscillate about the point or return back to it. It takes just a little mathematics, but it can be shown that this particular point is stable and that a slight perturbation will cause the fractions to oscillate around the point but not to move away from it, case B in Figure 12.4.

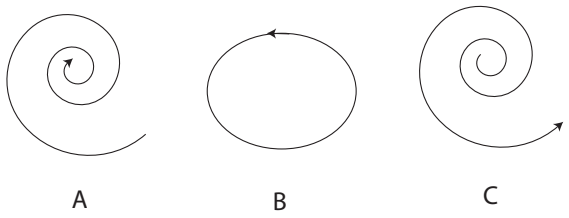


Figure 12.4: When a perturbation of some kind pushes the population off of an equilibrium point, the resultant dynamics can be stable as in cases A and B or unstable as in case C. It turns out that for Rock-Paper-Scissors the resultant trajectory will be stable in that it will oscillate around the equilibrium point.

Modeling a game as being played by a population of players, might seem odd for Rock, Paper Scissors, but if the rules are changed to another game you will see that the methodology can lead to important understandings. To that we will turn to a public goods game.

In this game workers can opt to either work or not. If they do they receive their wage minus an expenditure; if they do not work they do better receiving their wage. The wage is a based on the number of people who actually worked, divided by the number of workers (whether they worked or not). So you can see that this game has a version of the prisoner's dilemma in it in that it is attractive to 'defect,' or not work at all and still get something. However a variant makes this more interesting and that is to give workers the option of sitting out for a fixed sum. What this does is set up a cycle. Initially, workers find it profitable to sit out, as the remaining workers make enough of a sum to be profitably shared. But as the number of non-wokers increases, the total value of their goods produced dwindles to the point where it is better to sit out, which workers start to do. But as the total number of

workers in the wage pool decreases, then a person who chooses to work can do better than sitting out, so he or she will choose to work. Thus there is a repeating cycle.

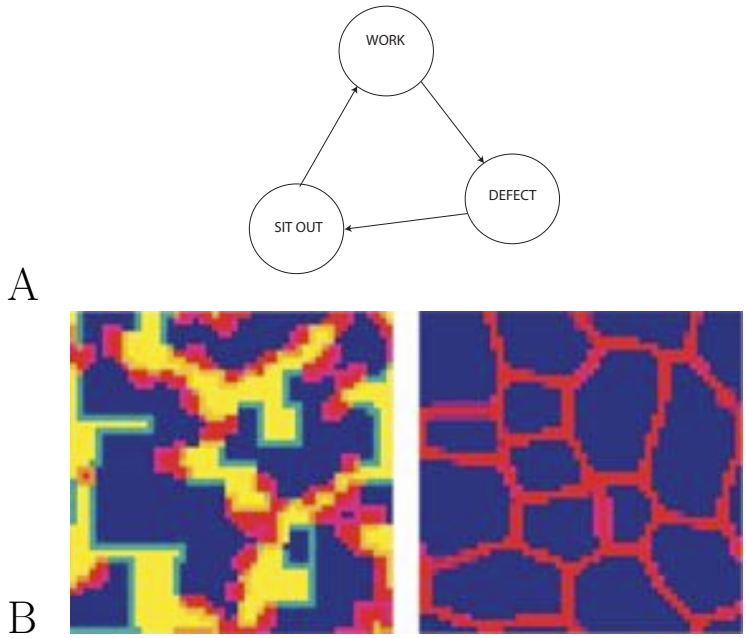


Figure 12.5: A) The representation of the cycle of choices made in a public goods game. B) The game can be played on a grid where players adopt the most succesful strategy in a $3x3$ grid neighborhood. Blue = Wokers, Red =Shirkers, Yellow= sitting Out. For certain values of the parameters the strategies exhibit a dynamics with waves moving accross the grid. A is a snapshot of such a case. For other values, Sit-outs die away and a stationary pattern of Workers and Shirkers is observed.[Permission Pending]

12.6 Summary and Key Points

The snail darters' IPD-like behavior is enormously revealing. It means that such behaviors are coded in the genes for those fish. But if its done for fish then it can be done that way for mammals too. What game theory provides us with is a framework for addressing this issue. With this framework the sensitivity of the strategies to various parameters can be characterized.

We also can speculate on how the embedding of game theory dynamics is done, given our perspective on emotions. As Rolls suggests, emotions can be seen as the hallmark of differences between what happened and what we expected to happen or, if the event is in the future, what might happen based on current evidence and what we think should happen. These differences are signaled as feelings. Thus it stands to reason that if we are genetically programmed to cooperate and we do not come through, there are going to be bad feelings. In fact we need to have the bad feeling to balance the delicate dopamine books. The math tells us that we are on average programmed to be good. Feelings are our experience of the running algorithms.

The hallmark work of Hamilton in showing the genes invest according to their fractional contribution is, from the standpoint of computation, a straightforward process. But when we turn to altruism between non relatives computation tells a different story. All the current game theory models, particularly when interpreted on an individual basis, show that the behaviors are delicate and are sensitive to the initial assumptions and payoffs. Here computation is again playing a fundamental role in laying bare the parameters of the human condition.

Bibliography

[1] Robert Axelrod. *The Evolution of Cooperation*. Basic Books, 1985.

[2] Richard Dawkins. *The Selfish Gene*. Oxford University Press, 1976.

[3] Paul Glimcher. *Decisions, Uncertainty and the Brain*. MIT Press, 2003.

[4] William Poundstone. *Prisoner's Dilemma*. Doubleday, 1992.