

## 16. Altruism: How can evolution promote group cooperation?

An important question for evolutionary theorists is: If I can only pass on my *own* genes, why should I ever do anything nice for you? Why might an organism *A* perform an action that increases the replicative fitness of some other organism *B*, even though *A*'s action does *not* increase *A*'s own replicative fitness?

In the 1960's-1970's, the idea of *kin-selection* arose to explain this: I help my brother because we share about a quarter of our genes, so by helping him, I promote the replication of a quarter of my own genes. By this accounting, I am twice as willing to save myself as to save my child; I am twice as willing to save my child as to save my sibling; and I am twice as willing to save my sibling as to save my half-sibling.

The important idea behind kin-selection is that selection does not reward individuals, but the replication of structures: replicating my genes is what contributes to population growth – not my own personal survival.

But if my fitness depends on your presence, it is in my own interest that I do what I can to ensure your presence. For example, kin-selection does not explain why African crows adopt and look after the babies of other crows, with whom they are not related. This example makes clear that selection rewards the replication of *developmental systems*. Susan Oyama used the term *developmental system* (DS) in 1986 to describe the complete system of genes, organisms and environmental conditions that are necessary for replication. To replicate, crows need crow genes, bodies, worms and seeds to eat, nests to sleep in, and other crows to help with all the work. By adopting young crows to share the work of survival, a gang of crows increases the replicative fitness of the gang and its entire DS, and it is the *DS* that stably survives.

So: how do organisms evolve to start working together to form a stable DS? Research in this area goes under the names of altruism, cooperation and group selection, and often focuses on a 2-person game called the *Prisoners' Dilemma*.

### The Prisoners' Dilemma

Biology is full of *mutualisms*: situations in which living systems cooperate to the benefit of each other. Your cells, nuclei and mitochondria work together to benefit the one sperm or egg that will form your children; ants and aphids provide each other with protection and honey-dew; the zooids in a Portuguese man-of-war are so dependent on each other that they are physiologically inseparable. Yet in each of these cases, it is possible for the participants to benefit by *defecting* from this cooperation: cells can grow into tumours, and ants could ignore aphid predators. How do such mutualisms arise, and why do participants usually not defect?

At each instant, two interacting organisms have two choices: Cooperate with the other or Defect. If they both try to Defect on each other, they might both benefit minimally, achieving an average payoff of 1 each. On the other hand, if they both Cooperate, they both gain the benefits of cooperation with a payoff of 4 each. The problem is, though, that if the aphid Cooperates and the ant Defects, the Defector gets a free meal of value 5, while the Co-operator gains absolutely nothing (0) from the encounter. The payoff matrix looks like this:

$$A_{PD} = \begin{matrix} & \begin{matrix} C & D \end{matrix} \\ \begin{matrix} C \\ D \end{matrix} & \begin{pmatrix} 4 & 0 \\ 5 & 1 \end{pmatrix} \end{matrix} = \begin{matrix} & \begin{matrix} C & D \end{matrix} \\ \begin{matrix} C \\ D \end{matrix} & \begin{pmatrix} R & S \\ T & P \end{pmatrix} \end{matrix} \quad (\text{Prisoners' Dilemma payoff})$$

This is the Prisoners' Dilemma (*PD*): Shall I risk cooperating, because that will reap benefits in the long-run, or shall I profit from the other's willingness to cooperate by defecting myself?

This is not the only possible form of the *PD* payoff matrix – any form satisfying the requirements  $T > R > P > S$  and  $R \geq (T + P)/2$  describes a *PD* encounter. Here,  $T$  is the temptation to defect, which is greater than the reward  $R$  for mutual cooperation.  $R$  is in turn greater than the punishment  $P$  for mutual defection, and worst of all is the sucker’s payoff  $S$  if I cooperate in the face of defection. The idea is that not cooperating, but *mutually* cooperating, is beneficial; if no-one else cooperates, then cooperating is bad for me.

- ? Verify that our payoff matrix  $A_{PD}$  represents the payoffs for the Prisoners’ Dilemma.

Now consider the following rational argument: If I cooperate, then you have a choice between gaining 4 points by cooperating or 5 points by defecting – so of course you defect. On the other hand, if I defect, you have a choice between gaining 0 points by cooperating or 1 point by defecting – so of course you defect. Thus, D *dominates* C: the second row of the payoff matrix shows that playing D *always* does better than playing C. If both players choose their move rationally, both will defect, so all they will ever gain is 1 point per encounter.

We can look at this dilemma in terms of a population of two strategy types: C’s have a frequency of  $x_1$  in the population and D’s have a frequency of  $x_2$ . In this case, C’s have an average payoff of  $r_1 = 4x_1$ , while D’s have an average payoff of  $r_2 = 5x_1 + x_2 = 4x_1 + 1$  (since  $x_2 = 1 - x_1$ ). But this means defectors always have a higher fitness than co-operators, and so will drive the co-operators to extinction. As a result, all individuals receive a far lower payoff than would be possible if they instead cooperated!

*This is the problem of the Commons: How can any population achieve the benefits of cooperating in their use of some common resource, if they as individuals always benefit from defecting?*

Cooperation can become beneficial in the *iterated PD*, where the same two players play the game  $m > 1$  times. Consider the following two strategies from Martin Nowak:

- **Grim:** cooperates in the first round, then cooperates if the other player doesn’t defect; if the other player defects just once, Grim never forgives, but will always defect.
- **AllD:** always defects, in all eternity, amen.
- ? Explain why this is the payoff matrix for these strategies:

$$\begin{array}{cc|cc} & & \text{Grim} & \text{AllD} \\ \text{Grim} & \left( \begin{array}{cc} mR & S + (m-1)P \\ T + (m-1)P & mP \end{array} \right) \\ \text{AllD} & & & \end{array}$$

If  $mR > T + (m-1)P$ , AllD does not dominate Grim. Indeed, Grim is also a *strict Nash equilibrium*: That is, two Grim players can never improve their score by switching to AllD. In evolutionary terms, if everyone in a population plays Grim, then a mutation that tries AllD can never invade the population, provided  $m$  is greater than the critical value  $(T - P)/(R - P)$ .

- ? Explain why AllD is also a strict Nash equilibrium, provided  $P > S$ .

So, if cooperation is once present, it can remain stable in the population; however, we do not yet understand how cooperation might evolve through mutation.

### Reactive strategies

In 1978, Robert Axelrod organised an iterated *PD* tournament for people from all over the world. Fourteen people submitted various *PD* strategies, and Axelrod used a computer to play all fourteen strategies against each other. The winner over all other strategies was Tit-For-Tat (TFT), submitted by the game theorist Anatol Rapoport. TFT is a forgiving version of Grim: it cooperates on the first round, then always plays whatever the other player played on the previous round of the iterated *PD*.

- ? Explain why the following matrix correctly describes the average payoff for TFT against AllD, where  $\bar{m}$  is the *average* number of rounds over which PD is iterated.:

$$\begin{array}{ccccc} & & \text{TFT} & \text{AllD} & \\ \text{TFT} & \left( \begin{array}{cc} \bar{m}R & S + (\bar{m} - 1)P \\ T + (\bar{m} - 1)P & \bar{m}P \end{array} \right) & & & \\ \text{AllD} & & & & \end{array}$$

Again, TFT is a strict Nash equilibrium – it can resist invasion by AllD, provided  $\bar{m} > (T - P)/(R - P)$ , but TFT has the advantage over Grim that it is not locked into defection: it can start cooperating again if the other player cooperates.

TFT is an example of a ***reactive strategy***: it decides what to do based on what happened on the previous iteration. We can define a reactive strategy  $S(p, q)$  in terms of two probabilities  $p$  (the probability that I will cooperate if you *cooperated* in the previous round) and  $q$  (probability that I will cooperate if you *defected* in the previous round).

- ? Explain why  $S(0,0)$  represents AllD, while  $S(1,0)$  denotes TFT.
- ? How would you describe a strategy with the representation  $S(1,0.2)$ ?

We can also describe the results of this round of PD in terms of four possible states (my move and your move): 1:CC, 2:CD, 3:DC and 4:DD. State 1 means we both cooperate; 2 means I cooperate but you defect; 3 means I defect and you cooperate; and 4 means we both defect.

Using these two ideas (( $p, q$ ) strategies and state-change), we can think of the dynamics of iterated PD as a *Markov chain*: a sequence of probabilistic transitions from the state (CC, CD, DC or DD) of one round to the state (CC, CD, DC or DD) of the next round. That is,  $\mathbf{x}_{t+1} = \mathbf{M} \cdot \mathbf{x}_t$ , where

$$\mathbf{M} = \begin{pmatrix} CC & CD & DC & DD \\ CC & p_1 p_2 & q_1 p_2 & p_1 q_2 & q_1 q_2 \\ CD & p_1 (1 - p_2) & q_1 (1 - p_2) & p_1 (1 - q_2) & q_1 (1 - q_2) \\ DC & (1 - p_1) p_2 & (1 - q_1) p_2 & (1 - p_1) q_2 & (1 - q_1) q_2 \\ DD & (1 - p_1) (1 - p_2) & (1 - q_1) (1 - p_2) & (1 - p_1) (1 - q_2) & (1 - q_1) (1 - q_2) \end{pmatrix}$$

- ? Verify that  $\mathbf{M}$  is a *stochastic matrix* – that is, the sum of entries in each column is 1. Why must this be true? (Hint: Think about how a basis vector is transformed by  $\mathbf{M}$ .)

When we discussed mutation matrices, we saw that they were stochastic matrices, and had an eigenvalue of 1 whose eigenvector  $\mathbf{x}^*$  is the long-term result of applying the stochastic matrix repeatedly over time. In fact, this property is quite generally true of all stochastic matrices, and so we can say with safety that in the long-term, the matrix  $\mathbf{M}$  will shift the population to a stable distribution  $\mathbf{x}^*$  satisfying an identity arising from the following eigenvalue equation:

$$\mathbf{x}^* \equiv \mathbf{M} \cdot \mathbf{x}^*$$

By substituting the above definition of  $\mathbf{M}$  into this eigenvalue equation, Martin Nowak was able to show that we can always calculate the *long-term payoff*  $A$  for any player that uses the strategy  $S_i(p_i, q_i)$  against strategy  $S_j(p_j, q_j)$  in this way:

$$(A)_{ij} \equiv A(S_i, S_j) = R \cdot s_{ij} s_{ji} + S \cdot s_{ij}(1 - s_{ji}) + T \cdot (1 - s_{ij})s_{ji} + P \cdot (1 - s_{ij})(1 - s_{ji}),$$

where  $\begin{cases} s_{ij} \equiv \frac{r_i q_j + q_i}{1 - r_i r_j} \\ r_i \equiv p_i - q_i \\ |r_i r_j| < 1 \end{cases}$  **(Nowak's reactive strategy payoff)**

Using these equations, we can calculate the long-term payoffs for any two strategies for which  $|(p_i - q_i)(p_j - q_j)| \neq 1$ . So, what do we do if  $p_i - q_i = \pm 1$ ?

- ? Explain why the only two strategies satisfying  $p_i - q_i = \pm 1$  are  $S(1,0)$  and  $S(0,1)$ . Find names for these two strategies.

Your answer to this learning exercise means there are two strategies (one of them extremely important!) whose long-term payoff we cannot calculate using Nowak's reactive strategy formula. There are two ways of dealing with this problem: either we specially calculate these cases individually, or we avoid them by using probabilities on the open interval  $p, q \in (0,1)$ . In the final project for this course, you should use the *second* of these two alternatives.