

Phage-lift for game theory

Martin A. Nowak and Karl Sigmund

The prisoner's dilemma is a classic of game theory in which acting for individual advantage is pitted against acting for collective benefit. An example has been identified among clones of a virus that infects bacteria.

A virus is a natural-born cheat that makes its living by exploiting the vital functions of a host cell. Small wonder, then, that viruses also exploit each other. By neatly combining theory and experiments, Turner and Chao (page 441 of this issue¹) have managed to demonstrate that certain phages — viruses that infect bacteria — actually engage in the prisoner's dilemma, that archetypal trap between cooperation and non-cooperation. Evolutionary game theorists will see this paper as a landmark. Indeed, it will be difficult to find players more primitive than the phage $\phi 6$ and its mutant clone $\phi H2$, stubby chunks of RNA that for their replication depend on a bacterial cell, and are therefore the subject of learned discussions as to whether they constitute life or not.

The prisoner's dilemma was devised by game theorists barely 50 years ago. Today, it seems difficult to conceive how moral philosophers, political thinkers or evolutionary biologists could ever have managed without it²⁻⁴. It is not much of a game, to be honest. Two players each have two options, to cooperate or not cooperate (defect). If both cooperate, they receive a reward, R , which is larger than the punishment, P , obtained if both defect. If one defects and the other cooperates, the defector obtains a payoff, T (the temptation), which is greater than R , and the cooperator receives the sucker's payoff S , which is less than P . So $T > R > P > S$. Because it pays more to defect, no matter whether the other cooperates or not, a rational player is bound to defect. Two rational players, therefore, will each end up with payoff P , instead of the reward R . Lowly $\phi 6$ usually does better and achieves the reward, so one might wonder whether rationality is really the gift it is supposed to be.

As the exemplar for the clash between individual advantage and collective benefit, the prisoner's dilemma was originally used to study the concept of rational choice, and to test actual human behaviour. In 1981, in a seminal paper by Axelrod and Hamilton⁵, it was applied to the evolution of cooperation in biological societies. Axelrod and Hamilton used computer simulations to display the emergence of cooperation in artificial populations, and suggested a wealth of biological examples, ranging from hairless primates

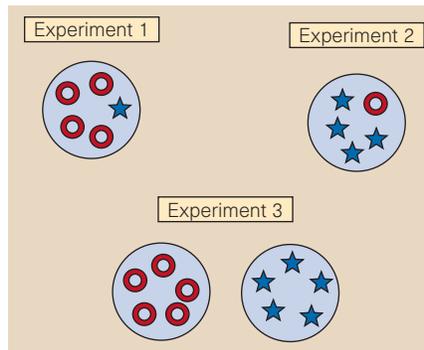


Figure 1 Bacteriophages $\phi 6$ (red) and $\phi H2$ (blue), and the prisoner's dilemma. First, Turner and Chao¹ measured the fitness of $\phi H2$ relative to that of $\phi 6$ in cells that contain mostly $\phi 6$. They thus determined $f(\phi H2, \phi 6)/f(\phi 6, \phi 6) = 1.99$. Second, they measured the same fitness ratio in cells that contain mostly $\phi H2$, and obtained $f(\phi H2, \phi H2)/f(\phi 6, \phi H2) = 1.28$. Third, they mixed cells that contained either $\phi 6$ or $\phi H2$ (but not both) and thereby measured the fitness of $\phi H2$ in cells that contain only $\phi H2$ relative to the fitness of $\phi 6$. They found $f(\phi H2, \phi H2)/f(\phi 6, \phi 6) = 0.83$. Because $f(\phi 6, \phi 6) = 1$ is the point of reference, Turner and Chao obtained the payoff values characteristic of the prisoner's dilemma, as described in the text.

engaged in trench warfare to bacteria living in their guts, in which the principles of the prisoner's dilemma might apply.

In the following years, both computer simulations and study of real-life occurrences of the prisoner's dilemma were expanding areas of research, but they did not grow at an equal pace. It proved much easier to do the simulations, and the empirical evidence lagged sadly behind. The same handful of examples were invoked time and again: vervet monkeys uttering alarm calls, sticklebacks and guppies engaged in predator inspection⁶, vampire bats feeding their hungry fellows. In most cases, the jury is still out on whether these are bona fide instances of the prisoner's dilemma⁷. The underlying problem is the bug-bear of evolutionary game theory: the currency for the payoff values is Darwinian fitness, which is notoriously difficult to measure for monkeys hiding in the bush, bats clustering in cave-roofs, and fish darting in and out of shoals.

With phages, the job becomes doable. The two strategies are embodied in the usual type of $\phi 6$ (the cooperator), and a mutant called $\phi H2$ (the defector) which manufactures fewer of the intracellular products needed for replication of the phages. Turner and Chao¹ measured the relative fitness of the two types in bacterial cultures by means of a genetic marker, cleverly exploiting the fact that the defectors' fitness is greater when they are rare. If the fitness of a $\phi 6$ phage in a $\phi 6$ -infested cell is set equal to 1, then that of a $\phi H2$ phage is almost double ($R = 1$ and $T = 1.99$) (Fig. 1). The fitness of a $\phi H2$ -defector in a $\phi H2$ -infested cell turns out to be $P = 0.83$, and that of a $\phi 6$ -phage in such a cell is $S = 0.65$. This is precisely the rank ordering required for the prisoner's dilemma.

This was by no means a foregone conclusion, even when one strategy is known to be more cooperative than the other. Suppose, for instance, that two cars are caught in a snowdrift⁸. To cooperate means getting out and starting shovelling. If the other driver does this, you can improve your own payoff by defecting (and staying in your heated car). So T is larger than R . But if the other player defects, you are well advised to get out and start shovelling. This is better than spending the night in the car. Hence S is larger than P , in contrast to the prisoner's dilemma.

In this case we end up with the payoff ranking of the so-called chicken game. In an evolutionary setting, defectors will not take over when playing this chicken game. They can invade a population of cooperators, but cooperators can also invade a population of defectors. The outcome is a mixed population. Examples of viral 'chicken' defectors have been known for a while. They quite literally have a defect that means that they cannot reproduce in the absence of complete viruses; in a population consisting entirely of their own kind, their fitness is zero. In this case, S is larger than $P = 0$, and we have a chicken game instead of a prisoner's dilemma.

It seems to us that such a possibility should also be tested carefully in interactions among animals with a cognitive apparatus: in such cases it is likely that a social norm will evolve which determines which of the two players does the cooperating and which does not. For instance, if two lionesses are jointly engaged in territorial defence of their pride, one may consistently be bolder than the other in carrying out defensive duties⁹. If the cost of losing the territory is higher than the risk of an injury, this would be precisely what the theory predicts.

Game theory purists may still quibble that Turner and Chao's matrix does not satisfy the condition $2R > T + S$. This condition is usually added to the definition of the prisoner's dilemma to rule out the possibility that one of the two players cooperates, the other defects, and both then share their total payoff. With phages, this is not to be feared

because it exceeds their capabilities. Another technical objection is that the phages crowded in the bacterial cell are not engaged in a pairwise contest, the usual context of the prisoner's dilemma. Again, this is no serious matter. What counts is that, whatever the ratio of defecting to cooperating phages in a cell, defectors always have an advantage. Necessarily, they will multiply faster, to the detriment of the average fitness.

So why haven't they taken over? Why is the predominant form the good, helpful $\phi 6$ rather than its lazy cousin $\phi H2$? We know little about the ecology of $\phi 6$ (even its natural bacterial host is unknown). But the origin of the defector gives us a clue, for the mutant $\phi H2$ evolves only if the multiplicity of infection is high¹⁰; that is, if there are many phages invading each host cell. In such circumstances, a phage is likely to find itself in a host cell together with phages from another clone, and then it pays to exploit them. In contrast, if the multiplicity of infection is low, the potential suckers are probably closest kin, members of the same clone, and exploitation would be ultimately self-defeating.

It should be noted, however, that the life cycle of phage $\phi 6$ (reproducing in bacteria which eventually burst, and reentering new bacteria) marks it as an ideal candidate for a specific type of group selection¹¹ — which can, more orthodoxly, be viewed as individual-based selection for the ability to build successful groups¹². Is cooperation due to being related or to being in the same boat? Even non-related phages may be better off by cooperating, and producing their full share of intracellular products, if the defective $\phi H2$ has an edge in the competition within the cellular compartment only, and not over the full life cycle. Exploring this issue calls

for further experiments with artificially increased competition based on high multiplicity of infection: these phages may become a testing ground for arguments on kin selection versus group selection, or on the evolution of virulence¹³, just as they have been used for studying sexual recombination¹⁰.

Finally, we recall that the role of compartments in sheltering cooperators from being exploited is crucial to certain hypotheses on the origin of life^{14,15}. Today, phages are arguably the simplest players of the prisoner's dilemma. But it is conceivable that, a few billion years ago, still more primitive molecules were engaged in that game. □

Martin A. Nowak is at the Institute for Advanced Study, Princeton, New Jersey 08540, USA.

e-mail: nowak@ias.edu

Karl Sigmund is at the International Institute for Applied Systems Analysis, Laxenburg, Austria, and the Institute of Mathematics, University of Vienna, Strudlhofgasse 4, A-1090 Vienna, Austria.

e-mail: ksigmund@esi.ac.at

1. Turner, P. E. & Chao, L. *Nature* **398**, 441–443 (1999).
2. Trivers, R. Q. *Rev. Biol.* **46**, 35–57 (1971).
3. Axelrod, R. *The Evolution of Cooperation* (Penguin, London, 1989).
4. Sigmund, K. *Games of Life* (Penguin, London, 1995).
5. Axelrod, R. & Hamilton, W. D. *Science* **211**, 1390–1396 (1981).
6. Milinski, M. *Nature* **325**, 433–435 (1987).
7. Dugatkin, L. A. *Cooperation Among Animals: An Evolutionary Perspective* (Oxford Univ. Press, 1997).
8. Sugden, R. *The Economics of Rights, Cooperation and Welfare* (Blackwell, Oxford, 1986).
9. Heinsohn, R. & Packer, C. *Science* **269**, 1260–1263 (1995).
10. Turner, P. E. & Chao, L. *Genetics* **150**, 523–532 (1998).
11. Wilson, D. S. *Am. Nat.* **149**, 336–351 (1997).
12. Ridley, M. *The Origin of Virtue* (Penguin, London, 1996).
13. Nowak, M. A. & May, R. M. *Proc. R. Soc. Lond. B* **255**, 81–89 (1994).
14. Eigen, M. & Schuster, P. *The Hypercycle: A Principle of Natural Self-Organization* (Springer, Heidelberg, 1979).
15. Maynard Smith, J. & Szathmari, E. *The Origins of Life* (Oxford Univ. Press, 1999).

case in which there is reasonable evidence for beamed (rather than isotropic) emissions of γ -rays^{2,4,9}. And it is the first in which a radio afterglow appeared after just one day, only to disappear the next^{2,3}.

The reason these measurements are so exciting is that, although significant progress has been made in understanding the γ -ray and afterglow emission in terms of the standard fireball-shock model, many skeletons still lurk in the closet. The emission from γ -ray bursts is thought to occur when one or more shock waves form in a relativistically expanding fireball following a cataclysmic explosion. An intense debate continues on whether the initial γ -rays arise in internal shocks within the original outflow of nucleons from the fireball, or in an external shock as the outflow is decelerated by the surrounding medium. Furthermore, an understanding of the 'central engine' of the burst remains elusive, in part because γ -ray bursts emit most of their energy during the first hundred or so seconds, decaying afterwards with a power-law slope. Until now, only γ -ray and some X-ray information has been available for this crucial early period — prompt optical and X-ray measurements are needed to further constrain models of the radiation mechanism.

Another problem is the energy budget. For instance, if GRB990123 were emitting energy isotropically at the measured redshift^{2,6} distance of $z = 1.6$ (corresponding to when the Universe was only 30% of its current age), it would require 4×10^{47} J (equal to the energy contained in two solar rest masses). This would strain to breaking point any model based on a stellar progenitor, such as the collapse of a massive rotating star (a hypernova or collapsar), or a merger involving neutron stars or black holes. The measurement of an immediate optical afterglow¹ and possible evidence for beaming^{2,4,9} provides much needed clues in attempting to address these issues.

The early optical detection of GRB990123 was achieved with a robotic optical camera, called ROTSE, installed on a fast-response swivelling mount. ROTSE consists of four telephoto lenses, which together can rotate and point more quickly (to reach any part of the sky within three seconds) than any large telescope. It responds to burst coordinates arriving by way of the Internet from the γ -ray coordinate network¹⁰, which in seconds transmits worldwide the rough locations (within a five arc-degree error) of all γ -ray bursts detected by NASA's Compton Gamma-Ray Observatory in space. With this instrument, Akerlof *et al.*¹ measured an optical brightness in excess of ninth apparent magnitude — which is too faint to see with the naked eye, but could be detected with a pair of binoculars — peaking 50 seconds after the trigger of GRB990123. In terms of brightness, this energetic burst is in the top

Gamma-ray astronomy

A burst caught in flagrante

Peter Mészáros

Gamma-ray bursts seem to delight in surprising astrophysicists. The most recent example is the event of 23 January 1999, analysed in three papers by Akerlof *et al.*¹, Kulkarni *et al.*² and Galama *et al.*³ in this issue of *Nature*, and in three related papers in *Science*^{4–6}. To appreciate the importance of these observations, we must remember that, for 23 years, γ -ray bursts were just that: brief pulses of γ -rays that pierced, for a fleeting instant, an otherwise pitch-black γ -ray sky. With rare exceptions, they left no trace at any other wavelengths, until in early 1997 the Italian–Dutch satellite Beppo–SAX succeeded in providing accurate X-ray measurements that allowed, after a delay of 4–6 hours, their position to be pinpointed and enabled follow-up measurements at optical and other wavelengths (for

example, ref. 7). This paved the way for the measurement of redshift distances, the identification of candidate host galaxies, and confirmation that they were indeed at cosmological rather than galactic distances (for example, ref. 8). This implies that γ -ray bursts are among the most powerful events in the Universe.

The burst known as GRB990123 is unique in at least four respects. It is the first to be detected optically while it was still emitting γ -rays — 22 seconds after being triggered¹. If GRB990123 emitted energy equally in all directions, then it would have to be the most energetic γ -ray burst detected so far^{2,6}. Alternatively, if the burst beamed its γ -rays in a jet, and an observer was near to the jet axis, then much less energy would be needed to produce the same intense flux. GRB990123 is the first