

On prisoners and cells

Karl Sigmund

In 1902, the Russian anarchist Kropotkin wrote a book on mutual aid in biological communities, feeling that Darwin's disciples tended to overlook its prevalence. This is no longer the case: altruism in a competitive world of replicators is now a favourite issue among darwinists, and the 'Prisoner's Dilemma' has emerged as the standard metaphor to model the conflict between mutual support and selfish exploitation. This model gains new perspective in two papers by M. Nowak and R. M. May, one on page 826 of this issue¹ and one due to appear in the *International Journal of Bifurcation and Chaos*².

In the Prisoner's Dilemma game, two players are engaged in a joint venture. Both have the options to cooperate or to defect (not cooperate). If both cooperate, they obtain more than if they both defect. But a player defecting unilaterally gains still more, while the cheated partner receives even less. Defecting is therefore the better choice, no matter what the other player does. If the payoff is reproductive success and strategies are inherited, defectors are bound to take over. Where does this leave mutual aid?

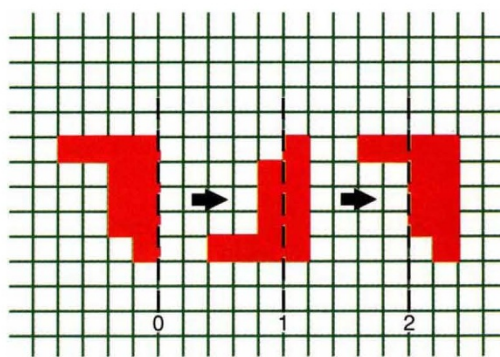
Repetition offers one way out of this impasse, as has been stressed by Axelrod and Hamilton^{3,4}. If retaliation is to be expected, it pays to refrain from defecting first because the bonus obtained by cheating in one round is offset by the risk of losing the partner's continued support. In order to retaliate in the iterated Prisoner's Dilemma, one has of course to recognize the other player. Defectors thrive in anonymous crowds, whereas mutual aid may be frequent between neighbours.

Roughly speaking, then, cooperating is good, but exploiting is better, as long as one can get away with it. Someone who is stuck in one place cannot get away with it, and therefore risks reprisals. But as Nowak and May point out, territoriality works even when no follow-up encounter is expected: it makes cooperation a viable option for the simple, single-shot Prisoner's Dilemma.

Indeed, imagine a population of players distributed on the squares of a chess board. Each player interacts only with the immediate neighbours. In the next generation, the square is inherited by whoever won most — neighbour or previous owner. Clearly, a lone cooperator will be exploited by the surrounding defectors. But four cooperators in a block can hold their own, because each interacts with more cooperators than a defector, as an 'outsider', can reach. If the bonus for cheating is not too large,

clusters of cooperators will grow. Of course, lone defectors will also do well: but, by prospering, they surround themselves with nasties and diminish their own returns.

Depending on the actual payoff values, the evolutionary dynamics can provide for various outcomes, but the usual result is regularly or irregularly fluctuating mosaics with both strategies holding on. In many cases, clusters of cooperators and clusters of defectors can both grow. If the initial conditions are random, computer simulations show an



Gliding on a wave of good will, a team of 11 cooperators moves through a world of defectors. Here, the pattern moves across a dotted line in three successive generations (time 0, 1 and 2).

endlessly milling spatio-temporal chaos, with metastatic tentacles flailing at each other. For symmetric initial conditions, a fantastic variety of patterns shows up, which could make a fortune for an enterprising tile manufacturer. That territoriality favours cooperation holds for many other neighbourhood geometries. It is likely to remain valid for random grids and stochastic transition rules, and for real-life communities; for example, the reciprocal altruism of aquatic cleaners and their obliging clients has been observed only among fish with fixed abode, not among those in the open sea.

Although Nowak and May are not the first to link loving one's neighbour with loving oneself (Matthew 22:29), they provide a new rationale for it, replacing the 'shadow of the future' (Axelrod's expression) by 'population viscosity' (a term due to Hamilton). Such a viscosity usually increases the probability that the coplayer is close kin, but Nowak and May make no assumptions about the genetic structure. The spread of successful strategies can be effected by cultural transmission just as well.

Spatial structures shoulder themselves into all kinds of population models nowadays, for instance, by stabilizing

host-parasite interactions or by protecting catalytic networks of replicating molecules; indeed, cooperation could be older than life itself. That apart, Nowak and May's combination of evolutionary game theory and cellular automata may well become a new fashion in computational mathematics. John von Neumann, who stood at the origin of both game and automata theory (and of the computer), would have enjoyed this union of his brainchildren.

Cellular automata⁵ provide a natural framework for evolutionary models with spatial structure. The best-known example is J. H. Conway's computer 'Game of Life'⁶. As discrete-time, discrete-space analogues of partial differential equations⁷, cellular automata are ideally suited for studying pattern formation and offer a vast field for mathematical exploration. The Nowak-May approach yields large classes of cellular automata, not just for the Prisoner's Dilemma but for other games such as 'Hawk-Dove'⁸ or 'Stone-Scissors-Paper'. Whereas Conway imposed arbitrary rules (after a lot of experimentation), the transition rules in the new automata emerge from evolutionarily relevant games in a natural way. Like the 'Game of Life', they exhibit glider structures (see figure) and a rich collection of other astonishing objects.

Few general principles are known for cellular automata, aside from a rough classification of their massively parallel dynamics of information transfer⁹, but it seems likely that some of Nowak and May's cellular automata are as complex as a universal computer. This would mean that, like 'Life', they contain patterns whose behaviour is on principle unpredictable. The possibility that the simplest conceivable two-player games can lead to fundamentally undecidable problems adds an intriguing dimension to Nowak and May's pulsating patterns of 'kaleidocooperation'. □

Karl Sigmund is in the Institute for Mathematics, University of Vienna, Strudlhofgasse 4, A-1090 Vienna, Austria.

- Nowak, M. A. & May, R. M. *Nature* **359**, 826–829 (1992).
- Nowak, M. A. & May, R. M. *Int. J. Bifurcation Chaos* (in the press).
- Axelrod, R. *The Evolution of Cooperation* (Basic Books, New York, 1984).
- Axelrod, R. & Hamilton, W. D. *Science* **211**, 1390–1396 (1981).
- Wolfram, S. *Theory and Applications of Cellular Automata* (World Scientific, Singapore, 1986).
- Berlekamp, E., Conway, J. H. & Guy, R. *Winning Ways for Your Mathematical Plays* (Academic, London, 1982).
- Murray, J. *Mathematical Biology* (Springer, New York, 1989).
- Maynard Smith, J. *Evolution and the Theory of Games* (Cambridge University Press, 1982).
- Langton, C. in *Artificial Life: Studies in the Science of Complexity* (eds Langton, C. et al.) 41–91 (Addison-Wesley, New York, 1992).