Equal Pay for All Prisoners

Maarten C. Boerlijst; Martin A. Nowak; Karl Sigmund

The American Mathematical Monthly, Vol. 104, No. 4 (Apr., 1997), 303-305.

Stable URL:
http://links.jstor.org/sici?sici=0002-9890%28199704%29104%3A4%3C303%3AEPFAP%3E2.0.CO%3B2-D

Equal Pay for All Prisoners

Maarten C. Boerlijst, Martin A. Nowak, and Karl Sigmund

By prisoners we mean, of course, players of the well-known Prisoner's Dilemma game (to be described presently). We shall show that there exist simple strategies for the infinitely iterated Prisoner's Dilemma that act as *equalizers* in the sense that all co-players receive the same payoff, no matter what their strategies are like.

The Prisoner's Dilemma game, a favorite with game theorists, social scientists, philosophers, and evolutionary biologists, displays the vulnerability of cooperation in a minimalistic model (see **[1]** to **[5]**). The two players engaged in this game can choose whether to cooperate or to defect. If both defect, they gain 1 point each; if both cooperate, they gain 3 points; but if one player defects and the other does not, then the defector receives 5 points and the other player only 0. The right move is obviously to defect, no matter what the other player does. As a result, both players earn 1 point instead of 3.

But if the same two players repeat the game very frequently, there exists no strategy that is best against all comers. The diversity of strategies is staggering. If we simulate on a computer populations of strategies evolving under a mutation-selection regime (with mutation introducing new strategies and selection weening out those with lowest payoff), we observe a rich variety of evolutionary histories frequently leading to cooperative regimes dominated by strategies like Pavlov (cooperate whenever the opponent's move, in the previous round, matched yours) or Generous Tit For Tat (always reciprocate your opponent's cooperative move, but reciprocate only two-thirds of the defections). Remarkably, all strategies of the iterated Prisoner's Dilemma, which can be very complex and make up a huge set, obtain the same payoff against some rather simple equalizer strategies.

More generally, let us consider a two-player game where both players have the same two strategies and the same payoff matrix. We denote the first strategy (row 1) by **C** (for 'cooperate') and the second (row 2) by **D** (for 'defect') and write the payoff matrix as

$$\begin{array}{ccc|cc} & & & \multicolumn{2}{c}{\text{Opponent}} \\ & & & \mathbf{C} & \mathbf{D} \\ \hline \text{You} & & \mathbf{C} & R,R & S,T \\ & & \mathbf{D} & T,S & P,P \end{array}. \qquad (1)$$

Such games include the Prisoner's Dilemma, where $T > R > P > S$, and the Chicken game, where $T > R > S > P$. (In the Prisoner's Dilemma case, R stands for the *reward* for mutual cooperation, P is the *penalty* for mutual defection, T is the *temptation* payoff for unilaterally defecting and S the *sucker* payoff for being exploited.)

Let us assume that the game is repeated infinitely often. A *strategy* in such a supergame is a program telling the player in each round whether to play **C** or **D**. The program may be history-dependent and stochastic: it specifies at every step the probability for playing **C**, depending on what happened so far. If A_n is the payoff in the n-th round, the expected long-run average payoff for a player is given by

$$\lim_{N\to\infty} \frac{A_1 + \cdots + A_N}{N} \tag{2}$$

provided it exists. It need not always exist: think of two players cooperating in the first 10 rounds, defecting in the next 100 rounds, then cooperating in the following 1000 rounds, etc.

Memory-one strategies are particularly simple. Such a strategy is given by the probability to play **C** in the first round, and a quadruple $\mathbf{p} = (p_R, p_S, p_T, p_P)$, where p_i denotes the probability that the player plays **C** after having experienced outcome $i \in \{R, S, T, P\}$ in the previous round. Some of the most successful strategies belong to this class, including Generous Tit For Tat $(1, 1/3, 1, 1/3)$ and Pavlov $(1, 0, 0, 1)$.

Theorem. *If* $\max(S, P) < \min(R, T)$, *then there exist, for every value* π *between these numbers, memory-one strategies* $\mathbf{p}$ *such that every opponent obtains the long-run average payoff* π *against a player using such a strategy. The vector* $\mathbf{p}$ *is given by*

$$(1 - (R - \pi)a, 1 - (T - \pi)a, (\pi - S)a, (\pi - P)a) \tag{3}$$

where a *is any real number such that* $1/a \geq \max(T - \pi, R - \pi, \pi - S, \pi - P)$.

Proof: The condition on a guarantees that the p_i are probabilities. Let us denote by $q_i(n)$ the conditional probability that the opponent plays **C** in the following round, given that the n-th round resulted in outcome i, and by $s_i(n)$ the probability that the outcome in the n-th round is i. By conditioning on round n, we obtain:

$$\begin{aligned} s_R(n+1) &= s_R(n)q_R(n)[1 - (R - \pi)a] + s_S(n)q_S(n)[1 - (T - \pi)a] \\ &\quad + s_T(n)q_T(n)(\pi - S)a + s_P(n)q_P(n)(\pi - P)a. \end{aligned} \tag{4}$$

Similarly,

$$\begin{aligned} s_S(n+1) &= s_R(n)(1 - q_R(n))[1 - (R - \pi)a] \\ &\quad + s_S(n)(1 - q_S(n))[1 - (T - \pi)a] \\ &\quad + s_T(n)(1 - q_T(n))(\pi - S)a + s_P(n)(1 - q_P(n))(\pi - P)a. \end{aligned} \tag{5}$$

Summing (4) and (5) yields the probability that you play **C** in round $n + 1$

$$\begin{aligned} s_R(n+1) + s_S(n+1) &= s_R(n)[1 - (R - \pi)a] + s_S(n)[1 - (T - \pi)a] \\ &\quad + s_T(n)(\pi - S)a + s_P(n)(\pi - P)a. \end{aligned}$$

Hence

$$\begin{gathered} a^{-1}[s_R(n) + s_S(n) - s_R(n+1) - s_S(n+1)] = \\ Rs_R(n) + Ss_T(n) + Ts_S(n) + Ps_P(n) - \pi[s_R(n) + s_S(n) + s_T(n) + s_P(n)]. \end{gathered} \tag{6}$$

Since the $s_i(n)$ sum up to 1, the right-hand side is just $A_n - \pi$, where A_n is the *opponent's* payoff in the n-th round (we must bear in mind that one player's outcome S is the other player's outcome T). Summing up (6) for $n = 1, \ldots, N$ and dividing by N, we obtain

$$\frac{1}{aN}[s_R(1) + s_S(1) - s_R(N+1) - s_S(N+1)] = \frac{A_1 + \cdots + A_N}{N} - \pi,$$

and hence

$$\lim_{N\to\infty} \frac{A_1 + \cdots + A_N}{N} = \pi. \qquad \blacksquare$$

A few final remarks. Two players using equalizer strategies are in *Nash equilibrium*, which means that neither has an incentive to change strategy. Nash equilibria exist for every game; for iterated games, they abound. Indeed, the so-called Folk Theorem in game theory states that every feasible pair of payoff-values exceeding the *minimax* (the highest payoff that a player can enforce, which in our case is $\max(S, P)$) can be realized by a Nash-equilibrium pair [**2**, p. 373]. Our theorem is related to this: the strategies are equalizers with memory one. Two players using such strategies have no reason to switch unilaterally to another strategy, since they cannot improve their payoff; however, they have no reason *not* to adopt another strategy either, since they will not be penalised. Since their opponent plays an equalizer strategy, they can switch to *any* other strategy, and not be worse off. If *both* players opt for a change, however, they are likely to end up in a non-equilibrium situation.

If a is chosen small enough, the runs of consecutive defections or cooperations can be made arbitrarily long. The condition $\min(R, T) > \max(S, P)$ and its converse are not only sufficient, but also necessary for the existence of such equalizer strategies. It is easy to construct other equalizer strategies. For example, play **C** until the opponent's mean payoff is larger than π, then play **D** until it is smaller than π, then play **C** until it is larger again, etc. However, such a strategy requires monitoring the opponent's entire payoff sequence. The point is that even within memory-one strategies, equalizers exist.

ACKNOWLEDGMENTS. Financial support from the Wellcome Trust (MAN) and the Austrian Forschungsförderungsfonds (KS) is gratefully acknowledged.

REFERENCES

1. R. Axelrod, *The Evolution of Cooperation*, Penguin, Harmondsworth, 1990.
2. K. G. Binmore, *Fun and Games: a Text on Game Theory*, Heath and Co, Lexington, Massachusetts, 1992.
3. M. Mesterton-Gibbons, *An Introduction to Game-theoretic Modelling*, Addison-Wesley, Redwood City, California, 1992.
4. M. A. Nowak, R. M. May, and K. Sigmund, The Arithmetics of Mutual Aid, *Scientific American* **272** (1995), 76–81.
5. K. Sigmund, *Games of Life: Explorations in Ecology, Evolution and Behaviour*, Penguin, Harmondsworth, 1995.

Maarten C. Boerlijst and Martin A. Nowak
Department of Zoology
University of Oxford
South Parks Road, Oxford OX1 3PS
UK
mcb@einstein.zoo.ox.ac.uk
nowak@nowak.zoo.ox.ac.uk

Karl Sigmund
Institut für Mathematik
Universität Wien
Strudlhofgasse 4, A-1090 Vienna
Austria
ksigmund@esi.ac.at