

Browning, L., & Colman, A. M. (2004). Evolution of coordinated alternating reciprocity in repeated dyadic games. *Journal of Theoretical Biology*, 229 549-557.

Evolution of coordinated alternating reciprocity in repeated dyadic games

Lindsay Browning^{a,*}, Andrew M. Colman^b

^a*Department of Experimental Psychology, South Parks Road, Oxford OX1 3UD, UK*

^b*School of Psychology, University of Leicester, Leicester LE1 7RH, UK*

Abstract

A genetic algorithm incorporating mutation and crossing-over was used to investigate the evolution of social behaviour in repeated Prisoner's Dilemma, Chicken (Hawk-Dove), Battle of the Sexes, and Leader games. The results show that the strategic structure of an interaction has a crucial determining effect on the type of social behaviour that evolves. In particular, simulations using repeated Prisoner's Dilemma and Chicken (Hawk-Dove) games lead to the emergence of genes coding for symmetric reciprocity and the evolution of mutual cooperation, whereas simulations using repeated Battle of the Sexes and Leader games lead to near-fixation of genes coding for asymmetric strategic choices and the evolution of coordinated alternating reciprocity. A mechanism is suggested whereby, in games with asymmetric equilibrium points, coordinated alternating reciprocity might evolve without insight or communication between players.

1. Introduction

The introduction of genetic algorithms (Goldberg, 1989; Holland, 1975; Riolo, 1992) enabled researchers to investigate the natural selection of social behaviour using simulations incorporating mutation and crossing-over. This research has focused almost exclusively on the repeated Prisoner's Dilemma game (e.g., Axelrod, 1987; Axelrod and Dion, 1988; Bendor *et al.*, 1991; Boerlijst *et al.*, 1997; Kraines and Kraines, 1989, 1995; Lindgren, 1991; Nowak and May, 1992; Nowak *et al.*, 1995; Nowak and Sigmund, 1992, 1993, 1998; Wu and Axelrod, 1995). Among the few who have studied the evolution of behaviour in other games are Binmore and Samuelson (1992); Crowley (2001); Friedman (1996); Lipman (1986); Posch (1999); and Roth and Erev (1995). However, even among the simplest strategic games, namely dyadic (two-player) games in which each player has just two strategies, there are 12 ordinally distinct, symmetric 2×2 games (Rapoport and Guyer, 1966).

Rapoport (1967) showed that, after excluding games that are strategically trivial in the sense of having equilibrium points that are uniquely Pareto-efficient, there remain four archetypal 2×2 games: Prisoner's Dilemma, Chicken (Hawk-Dove), Battle of the Sexes, and Leader, but there is no reason to suppose that the Prisoner's Dilemma game is more characteristic of naturally occurring strategic interaction than any other.

2. Strategic Structures

The findings reported in this article are based on simulations using the four archetypal games, with the payoffs shown in Fig. 1. The Prisoner's Dilemma game (Fig. 1A), defined by the inequalities $T > R > P > S$ (Fig. 1E), has a unique equilibrium point at (D, D) , corresponding to joint defection, and the D strategy is dominant for both players, in the sense that it yields a higher payoff than the C strategy irrespective of the co-player's choice, but each player receives a higher payoff if both choose C than if both choose D . The game was discovered at the RAND Corporation in 1950 and popularized by Luce and Raiffa (1957, pp. 94-102). A typical interaction in which the game is often believed to arise in nature involves

a pair of conspecifics that can either groom each other (*C*), to their mutual advantage, or defect (*D*) by declining to groom each other. The outcome is better for each player if both cooperate than if both defect, but a player receives the best possible payoff by defecting while the co-player cooperates, in which case the cooperative player receives the worst payoff, because the cooperator expends time and energy grooming but receives nothing in return.

		II		
		C	D	
I	C	3, 3	0, 5	I
	D	5, 0	1, 1	
		(A) Prisoner's Dilemma		

		II		
		C	D	
I	C	3, 3	2, 4	I
	D	4, 2	1, 1	
		(B) Chicken		

		II		
		C	D	
I	C	2, 2	4, 3	I
	D	3, 4	1, 1	
		(C) Battle of the Sexes		

		II		
		C	D	
I	C	2, 2	3, 4	I
	D	4, 3	1, 1	
		(D) Leader		

		II		
		C	D	
I	C	R, R	S, T	I
	D	T, S	P, P	
		(E) Generalized Matrix		

Fig. 1. Payoff values for simulations using the four archetypal 2×2 games. Shaded cells are Nash equilibrium points

The other three games lack dominant strategies and have dual, asymmetric equilibrium points in pure strategies – we ignore mixed-strategy equilibria that become available only if players can choose probability distributions over their pure strategies, which requires the use of randomizing devices. Games with asymmetric equilibria, when iterated, turn out to have interesting evolutionary properties that are lacking in the iterated Prisoner's Dilemma game.

Chicken (Fig. 1B), defined by $T > R > S > P$, is the prototypic dangerous game, because a player has to risk the lowest payoff to have a chance of the highest, and it has equilibrium points at (*D*, *C*) and (*C*, *D*), Player I preferring (*D*, *C*) and Player II (*C*, *D*). This game was first described by Russell (1959, p. 30). The classic biological example, introduced by Maynard Smith and Price (1973) and later named the *Hawk-Dove game* (Maynard Smith, 1982), occurs when a pair of organisms contesting a resource can engage in conventional fighting (*C*) or can defect by engaging in escalated fighting (*D*). Under reasonable assumptions, the outcome is best for a player that defects while the co-player cooperates; second-best for each if both cooperate; third-best for a player that cooperates while the co-player defects; and worst for each if both defect, because then the fighting is both escalated and protracted.

Battle of the Sexes (Fig. 1C), defined by $S > T > R > P$, also has equilibrium points at (D, C) and (C, D) , but in this game Player I prefers (C, D) and Player II (D, C) . The game was introduced by Luce and Raiffa (1957, p. 91). A typical biological interaction with this strategic structure might occur if two predators feeding on a kill are being harassed by scavengers. (Although we continue to use the symbols C and D for convenience, their interpretation in terms of cooperation and defection are hardly applicable to this game or to the game of Leader below.) Each can either ignore the scavengers (C), or temporarily abandon the kill to chase the scavengers (D). We may assume that the payoff is best for a player who ignores the scavengers and continues feeding while the co-player chases the scavengers; second-best for a player who unilaterally stops feeding to chase the scavengers and thus loses a little feeding time; third-best for each player if both ignore the scavengers and therefore lose some of the kill; and worst for each if both simultaneously abandon the kill to chase the scavengers, because in that case they run the risk of losing the whole kill.

Leader (Fig. 1D), defined by $T > S > R > P$, has equilibrium points at (D, C) and (C, D) and, as in Chicken, Player I prefers (D, C) and Player II (C, D) . This game was first identified and named by Rapoport (1967). It might occur in nature if two organisms need to escape from a predator through an escape route that is wide enough for only one of them at a time. Each player can either bolt for the opening (D) or wait for the co-player to bolt, with the intention of following immediately after (C). We may assume that the chances of escape are best for a player who bolts while the co-player waits; second-best for a player who waits and then follows after the co-player bolts; third-best for players who both wait for each other and thus fail to bolt; and worst for players who bolt simultaneously and thus run the risk of injuring each other or getting stuck, thus becoming easy prey.

In the repeated Prisoner's Dilemma game, the simple reciprocal strategy Tit for Tat (TFT) cooperates on round 1, then on every subsequent round t copies the co-player's choice from round $t - 1$. The strategy pair (TFT, TFT) is an equilibrium point (Binmore, 1992, pp. 368-369), corresponding to the (C, C) outcome, and so is (ALL D , ALL D), corresponding to the (D, D) outcome, which is the unique equilibrium point of the one-shot stage game. Mathematically, (TFT, TFT) corresponds to the point $(3, 3)$ and (ALL D , ALL D) to the point $(1, 1)$ in the basic outcomes of the one-shot Prisoner's Dilemma stage game.

Mathematically, the average payoffs per round that are possible in an indefinitely repeated Prisoner's Dilemma game lie in a *cooperative payoff region* defined by the convex hull of points representing the four outcomes of the one-shot game, namely $(3, 3)$, $(0, 5)$, $(5, 0)$, and $(1, 1)$. For example, if Player I adopts an unconditionally cooperative strategy and Player II simply alternates between cooperation and defection, then it is clear from an examination of Fig. 1A that Player I earns an average payoff of 1.5 per round, mid-way between R and S , while Player II earns an average payoff of 4, mid-way between R and T , and the point $(1.5, 4)$ lies in the interior of the cooperative payoff region. The so-called *folk theorem* of indefinitely repeated games establishes that equilibrium points are densely scattered throughout the part of the cooperative payoff region where the players' payoffs are at least as high as the payoffs they can guarantee for themselves by choosing a best reply to each of the co-player's strategies (for a straightforward proof, see Binmore, 1992, pp. 373-377).

An evolutionarily stable strategy or ESS (Maynard Smith, 1982; Maynard Smith and Price, 1973) is a strategy with the property that if most members of the population adopt it, no alternative strategy has a higher Darwinian fitness, and therefore none can invade the population by natural selection. Suppose a population consists mostly of individuals adopting some arbitrary strategy I , but a small fraction p of mutants adopt strategy J . An individual adopting strategy I receives a payoff of $E(I, I)$ with probability $1 - p$ and a payoff of $E(I, J)$ with probability p . An individual adopting strategy J receives a payoff of $E(J, I)$ with probability $1 - p$ and $E(J, J)$ with probability p . If the Darwinian fitness of each member of

the population before a series of contests is C , then after the contests the fitness of an individual adopting Strategy I , denoted by $W(I)$, is

$$W(I) = C + (1 - p)E(I, I) + pE(I, J),$$

and the fitness of an individual adopting strategy J , denoted by $W(J)$, is

$$W(J) = C + (1 - p)E(J, I) + pE(J, J).$$

If I is an ESS, then $W(I) > W(J)$ by definition. Because p is assumed to be small, either

$$E(I, I) > E(J, I), \tag{1}$$

or

$$E(I, I) = E(J, I) \text{ and } E(I, J) > E(J, J). \tag{2}$$

Conditions (1) and (2) are Maynard Smith's (1982) definition of an ESS, and it is the strategy that should normally be found in nature.

An ESS is invariably an equilibrium point, but not every equilibrium point is an ESS (Hofbauer and Sigmund, 1998, pp. 62-65; Samuelson, 1998, pp. 38-40). Neither (TFT, TFT) nor (ALL D , ALL D) is an ESS (Lorberbaum, 1994). The PAVLOV (Win-Stay, Lose-Change) strategy (Kraines and Kraines, 1989, 1995; Nowak and Sigmund, 1993), cooperates on round 1, then repeats its own previous-round strategy whenever its payoff is sufficiently high (usually 3 or 5 in Fig. 1A), or switches strategies otherwise. The strategy pair (PAVLOV, PAVLOV), corresponding to (C, C) , is an equilibrium point but not an ESS.

In repeated Chicken (Fig. 1B), strategies corresponding to the stage game's (C, D) and (D, C) equilibrium points remain in equilibrium, together with (TFT, TFT) and (PAVLOV, PAVLOV), both corresponding to (C, C) . In repeated Battle of the Sexes and Leader, strategies corresponding to the stage game's (C, D) and (D, C) equilibrium points remain in equilibrium, but (TFT, TFT) and (PAVLOV, PAVLOV) are out of equilibrium and are therefore not ESSs. In Battle of the Sexes and Leader, *coordinated alternating reciprocity* is more profitable to each player, yielding the payoff stream ..., (4, 3), (3, 4), (4, 3), ..., than joint cooperation, yielding ..., (2, 2), (2, 2), (2, 2), ..., and this raises interesting evolutionary problems, whereas in the version of Chicken used in this research, the payoffs from joint cooperation are as good as those from coordinated alternating reciprocity, because $2R = S + T$.

3. Genetic Algorithm

Axelrod (1987) ran his original genetic algorithm over 50 generations, and TFT-like strategies evolved; but a later Prisoner's Dilemma study (Lindgren, 1991), using different primordial strategies, reported alternating periods of stable reciprocity and instability. Several researchers have investigated patterns that emerge only after many generations (Kraines and Kraines, 1995; Leimar, 1997; Lorberbaum, 1994; Nowak and May, 1992; Nowak *et al.*, 1995; Posch, 1999). Wu and Axelrod (1995) found TFT to be evolutionarily successful in a 2000-generation replication incorporating noise. Noise has a radical effect on the course of evolution. See, for example, Kraines and Kraines (1995), Nowak *et al.* (1995), Posch (1999), and Wu and Axelrod (1995). The simulation described below was run over 1000 generations, with noise modelled via random mutations, and evolving strategies were paired against one another, as occurs in dyadic strategic interactions in nature.

The payoff values for the Prisoner's Dilemma game have been used by Axelrod (1987) and many other researchers. For the games of Chicken, Battle of the Sexes, and Leader,

payoffs were taken from Colman (1995). The procedure followed Axelrod (1987) and Wu and Axelrod (1995). Each simulation was started with 20 random strategies, a strategy being represented by a 70-digit binary string. The total number of strategies that can be represented in this form is $2^{70} \approx 10^{21}$. Each strategy was essentially a Moore machine designed to generate a move on the basis of any possible three-move history. For each outcome, a player receives one of four payoffs labelled *T*, *R*, *S*, or *P* in Fig. 1E, thus there are $4^3 = 64$ different three-move histories. A string of 64 binary digits therefore suffices to specify a choice for every three-move history, with 0 denoting a *D* choice and 1 a *C* choice. For example, if the digit associated with a three-move history *RRR* is 1, then after three consecutive *R* payoffs – three consecutive (*C*, *C*) outcomes – the strategy responds by playing *C*. A hypothetical three-move history was also necessary to get each game started. For this, six additional binary digits, named the *premise genes*, were added, making a total genome of 70 binary digits.

The offspring strategies that played in each subsequent generation were formed from the most successful strategies of the previous generation, using a genetic algorithm. The algorithm, designed along the lines of Holland (1975), implemented the following nine steps.

(1) The payoff values were assigned according to one of the four archetypal games shown in Fig. 1.

(2) The population size was set to x ($2 \leq x \leq 20$); the number of repetitions to be played by each pair of strategies in each generation to i ($i = 1, 2, \dots$); and the number of generations to g ($1 \leq g \leq 1000$). In the simulations reported here, $x = 20$, $i = 151$, $g = 1000$.

(3) An initial population was created by generating a pseudo-random 70-digit binary number for each of the x strategies.

(4) In each generation, each of the x strategies was paired with each of the others for i repetitions, until every strategy had played i repetitions with every other strategy in the population. Payoffs were assigned according to the values shown in the corresponding payoff matrix (Fig. 1) and were averaged over the i repetitions of each pair.

(5) At the end of each generation, after each strategy had played i repetitions with each of the others, each strategy's mean payoff over its $x - 1$ pairwise interactions was computed, and it was assigned a mating probability proportional to its score.

(6) For each offspring strategy, two strategies were randomly selected as parents, selection being proportional to mating probability scores. Each new generation consisted exclusively of the x offspring strategies.

(7) Genetic crossing-over was simulated by randomly selecting a number from 3 through 67 as the crossover point and breaking the 70-digit parent strategy's binary code at this point. The code before the crossover point from parent *A* and after the crossover point from parent *B* combined to form the offspring strategy for the next generation.

(8) For each offspring strategy, genetic mutation was simulated by randomly flipping one of the 70 binary digits of its genome from 0 to 1 or from 1 to 0. This mutation rate (1.4 per cent), artificially high in comparison with naturally occurring mutation, was chosen to facilitate rapid evolution of new strategies.

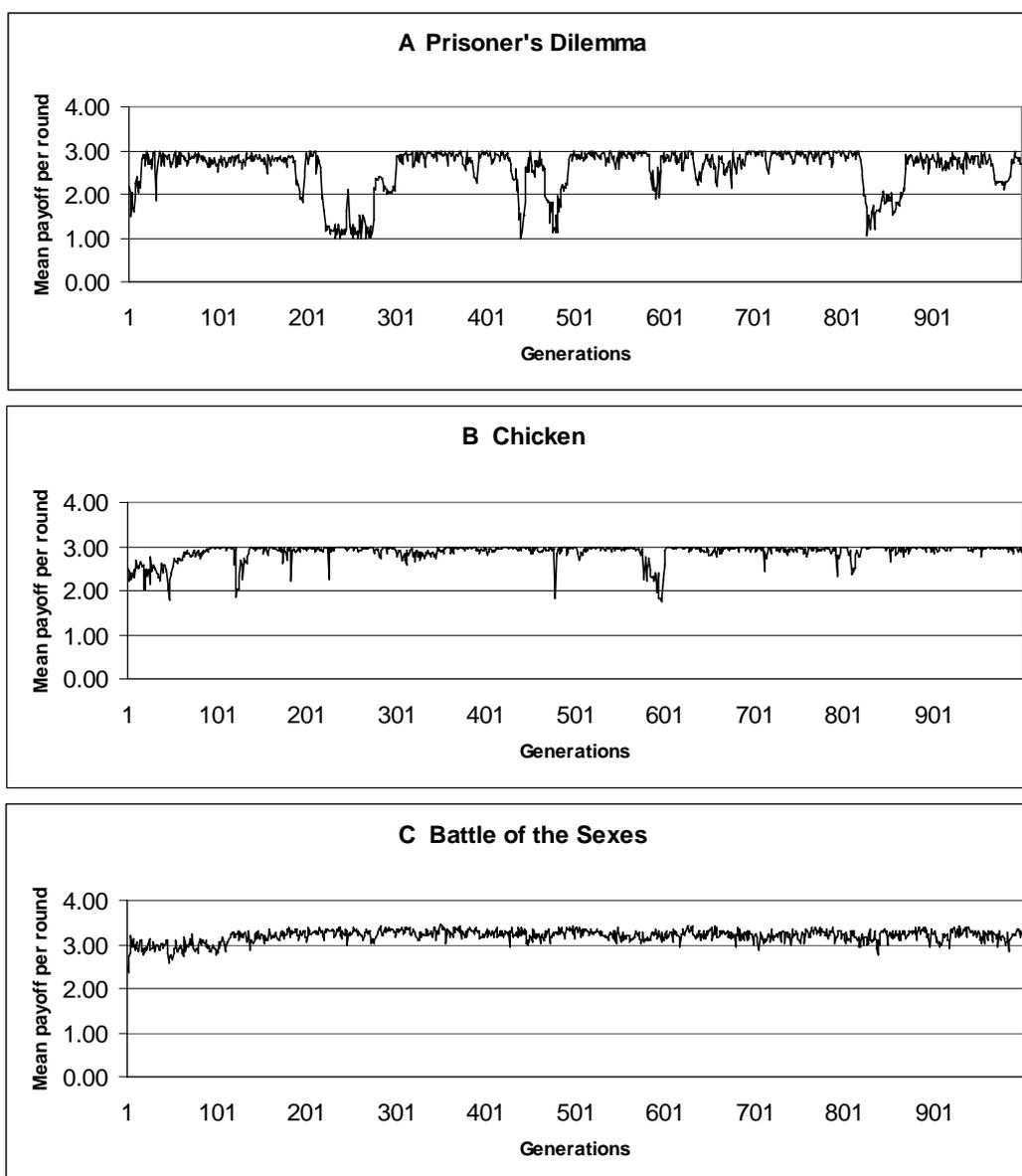
(9) Offspring strategies replaced parent strategies at each new generation, and the algorithm looped back to repeat steps 4 to 8 until g generations had been completed.

Each simulation used a population of 20 strategies, paired with each other for 151 repetitions (rounds). The simulation was continued for 1000 generations. The Prisoner's Dilemma simulation was replicated 10 times to check the reproducibility of the findings, and simulations of the other games were replicated five times.

4. Results

For each game, the pattern of results was very similar across replications. Fig. 2 therefore

displays mean payoffs per round for one typical simulation of each game structure, but the numerical data discussed below are averaged across replications. In the repeated Prisoner's Dilemma simulations (Fig. 2A), the mean payoff per round oscillated between equilibria close to 3.00 and 1.00, corresponding approximately to the payoffs for joint cooperation (C, C) and joint defection (D, D) respectively. In the repeated game, strategies leading to (D, D) outcomes, yielding payoffs of (1, 1), are in equilibrium, and TFT, leading to payoffs of (3, 3), corresponds to one of the many additional equilibrium points in the repeated game (Binmore, 1992, pp. 360-382). The data in Table 1 confirm across replications what is apparent by inspection of Fig. 2, that the jointly cooperative equilibrium tended to be more stable and persistent than joint defection. But it is clear that the Prisoner's Dilemma simulations did not stabilize fully, even after 1000 generations.



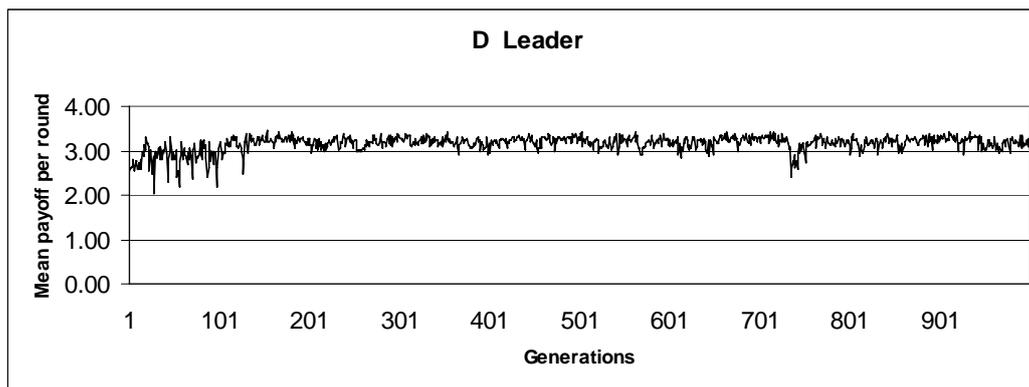


Fig. 2. Mean payoffs per round for 151 rounds over 1000 generations of representative examples of each of the four archetypal 2×2 games.

Table 1

Percentages of joint cooperative (C, C), asymmetric (C, D) and (D, C), and joint defecting (D, D) outcomes across 1000 generations of each simulation, averaged across replications

Game	(C, C)	(C, D)	(D, C)	(D, D)	Reciprocity
Prisoner's Dilemma	61.88	8.31	8.31	21.50	80.95
Chicken	66.72	12.90	12.90	7.48	87.60
Battle of the Sexes	10.32	41.76	41.76	6.16	88.60
Leader	9.17	40.90	40.90	9.13	81.65

Note: figures in the (C, D) and (D, C) columns are identical, because a (C, D) outcome for one player always corresponds to a (D, C) outcome for the co-player. The right-hand column shows, for the last 100 generations only, the reciprocity index, defined as the percentage of occasions on which players copied the strategy chosen by the co-player on the previous round, whether the co-player's strategy was C or D . The reciprocity index can be high not only when players repeatedly reciprocate co-players' cooperative strategy choices, as in sequences such as ..., (C, C), (C, C), (C, C), ..., but also in sequences of coordinated alternation such as ..., (C, D), (D, C), (C, D), ..., where it is also the case that each player copies the co-player's strategy from the previous round.

In the repeated Prisoner's Dilemma simulations, evolving strategies displayed some of the characteristics of TFT and PAVLOV. Perfect TFT or PAVLOV strategies would cooperate consistently against cooperative co-players, and many of the loci in their genetic algorithms would not be subject to selection pressures, because the corresponding three-move sequences would never occur. However, the sequence RRR (Fig. 1E), resulting from three successive joint cooperative outcomes, would be likely to occur repeatedly, and a TFT or PAVLOV strategy would respond with a cooperative C move, therefore the digital code at the RRR gene locus would invariably be 1 (coding for a C choice). During the last 100 generations, after 900 generations of evolution had occurred, the code at the RRR locus was 1 (coding for a C choice) in 79.78 per cent of population members. More directly suggestive of TFT is the value of the *reciprocity index* (Table 1), indicating a strong propensity of strategies to choose C in response to C or D in response to D .

None of the gene loci approached complete fixity and uniformity. A fixed gene, whether coded 0 (defect) or 1 (cooperate), should approach a relative frequency of 100 per cent in the population, although it cannot attain these values exactly because of variability arising from mutation in the genetic algorithm, resulting in approximately 1.4 per cent bit variability per strategy per generation. However, even during the last 100 generations, at no gene locus was the relative frequency of the 0 allele more than 76.34 per cent (PTP) or of the 1 allele more than 79.78 per cent (RRR).

For repeated Chicken (Fig. 2B), the mean payoff per round rose rapidly to around 3.00 (the expected payoff for joint cooperation) and remained close to that figure except for

relatively brief periods. The proportions of each of the four outcomes across 1000 generations, and the reciprocity index for the final 100 generations, are shown in Table 1. The equilibrium points of the stage game, (C, D) and (D, C) , are hard to achieve, because defection risks the lowest possible payoff. In the repeated game, joint cooperation, corresponding to payoffs of $(3, 3)$, is one of the new equilibria to emerge. During the last 100 generations, the *RRR* gene approached fixity, being coded 1 (cooperate) in 96.98 per cent of the population, and the reciprocity index was high (Table 1). No other gene approached fixity at either 0 or 1, the next nearest being the *SRP* locus, which was coded 0 (defect) in 87.44 per cent of the population.

For repeated Battle of the Sexes (Fig. 2C), the mean payoff rose over 120 generations to around 3.30 per round, not far short of the expected payoff from alternating (C, D) and (D, C) outcomes, and hovered there for the remaining 880 generations. The proportions of each of the four outcomes (Table 1) show the most frequent outcomes to have been the asymmetric (C, D) and (D, C) equilibrium points of the stage game. By the last 100 generations, fixation had tended to occur at just two gene loci, namely *TST* (virtually fixated at 97.94 per cent cooperation) and *STS* (99.08 defection). These data reveal the evolution of a form of asymmetric reciprocity through coordinated strategic alternation. Some evidence for this has already been reported by Crowley (2001), but Crowley focused on a different form of asymmetry, namely asymmetry in resource-holding power (RHP) or body size in Generalized Hawk-Dove games, and his data were based on strategies with only one-move memories. With certain payoff values, his classifier-system analysis provided suggestive evidence for the evolution of ‘complementarity’ in the form of alternating strategies, but these strategies did not evolve to high prevalence as in the present study, and he did not present any evidence for phase coordination between players.

In the Battle of the Sexes game reported in this article, the reciprocity index during the last 100 generations (Table 1) was high. It would be close to 100 per cent if the alternating strategies were invariably phase-coordinated, because in phase-coordinated alternation, each player simply repeats the co-player’s previous strategy choice. But it is reasonable to assume that a certain amount of manoeuvring was required in each dyadic interaction to lock phases into the mutually profitable (C, D) , (D, C) , (C, D) , ... cycle, and this may explain the shortfall in reciprocity.

Results for repeated Leader (Fig. 2D) show the mean payoff per round climbing to a relatively stable value close to 3.30, once again not far short of the payoff from alternating (C, D) and (D, C) outcomes. Over 1000 generations, the most frequent outcomes were (C, D) and (D, C) (Table 1), corresponding to the asymmetric equilibrium points of the stage game. By the last 100 generations, just one gene locus, *STS*, approached fixation at 93.94 per cent defection, and the next nearest to fixation was *TST* at 86.88 per cent cooperation. The reciprocity index was high, and the data are generally similar to those for the Battle of the Sexes, indicating once again the evolution of a form of coordinated alternating reciprocity.

Table 2 shows the percentages, at Generations 1, 500 and 1000, of three-move histories xyP , xyR , xyS , and xyT , where x and y stand for arbitrary payoffs P , R , S , or T in Fig. 1E. These figures show that, at the start, the percentages are evenly distributed in all four games, but after hundreds of generations of evolution, three-move histories terminating in R , representing mutually cooperative (C, C) outcomes, come to predominate in the Prisoner’s Dilemma and Chicken simulations, whereas three-move histories terminating in S or T , representing asymmetric (C, D) and (D, C) outcomes, come to predominate in the Battle of the Sexes and Leader simulations.

Table 2

Percentages of three-move histories xyP , xyR , xyS and xyT , where x and y stand for P , R , S or T (see Fig. 1E) at Generations 1, 500 and 1000

Game	Generation	xyP	xyR	xyS	xyT
Prisoner's Dilemma	1	28.61	22.83	24.30	24.26
	500	5.73	78.82	7.66	7.79
	1000	15.54	63.10	10.60	10.76
Chicken	1	26.32	26.50	23.58	23.60
	500	9.38	63.29	13.58	13.74
	1000	2.60	87.80	4.94	4.66
Battle of the Sexes	1	28.16	22.51	24.63	24.71
	500	4.94	6.89	44.07	44.09
	1000	4.14	7.79	44.11	43.96
Leader	1	28.58	22.35	24.52	24.54
	500	7.67	8.29	41.97	42.08
	1000	5.24	6.96	43.82	43.99

Supplementary analyses were performed for the Battle of the Sexes and Leader games, to remove any doubt that coordinated alternating reciprocity evolved in these two games. In five replications of the Battle of the Sexes simulation, the relative frequencies with which each of the 64 possible three-move histories occurred was counted directly in selected generations. The pattern of results was very similar in all five replications, and the following results are averages. In Generation 1, activity was distributed fairly evenly across the 64 three-move histories, and as evolution progressed over subsequent generations, certain three-move histories began to occur with greater relative frequency than others. In particular, in Generations 1, 10, 100, 300, 500, and 1000, the average percentages of the *STS* three-move history that occurred were 1.89, 9.51, 25.84, 33.12, 41.46, and 41.50 per cent respectively. The relative frequencies for *TST* were the same. This can be explained by the fact that *TST* always corresponds to a history of *STS* for the co-player. In Generation 1000, none of the other 64 possible three-move histories occurred with relative frequencies above chance level (1.56 per cent). Data from Generation 1000 show that *STS* and *TST* evolved to a joint relative frequency of 83 per cent of possible three-move histories. This confirms the data given above and indicates directly the relative frequency of coordinated alternating reciprocity that evolved in the Battle of the Sexes simulations. In the Leader game, the five replications led to very similar results once again, and the averaged relative frequencies of *STS* (and *TST*) in Generations 1, 10, 100, 300, 500, and 1000 were 2.53, 18.75, 30.20, 36.19, 35.71, and 40.28 per cent respectively. The only other three-move history that occurred above chance level in Generation 1000 was *RRR* (2.84 per cent). These findings suggest that coordinated alternating reciprocity in the Leader game evolved to more than 80 per cent of possible three-move histories in Generation 1000.

5. Discussion

The results show that the evolution of social behaviour depends in an interesting way on strategic properties of the underlying interactions. Evolution in the repeated Prisoner's Dilemma game, on which attention has overwhelmingly been focused in earlier research, differs radically from evolution in other archetypal games. In particular, the well known phenomenon of rapid evolution of symmetric cooperation through simple reciprocity was evident in the Prisoner's Dilemma and Chicken simulations, but in the other two games a more complex form of coordinated alternating reciprocity evolved.

In the Prisoner's Dilemma simulation, symmetric reciprocity evolved only imperfectly. After 900 generations of evolution, none of the gene loci approached fixation, although a majority of strategies had evolved a strong propensity to reciprocate, as would be predicted from both TFT and PAVLOV strategies. The gene locus for responding to a three-move history of joint cooperation (*RRR*) was coded for a cooperative response in 79.78 per cent of the population, far short of the almost invariable cooperation that would have been observed

had TFT or PAVLOV evolved perfectly. The reciprocity index also indicated a strong but by no means uniform TFT-like propensity to reciprocate cooperation and defection. The Chicken simulation resulted in greater symmetric reciprocity, with the *RRR* gene locus approaching cooperative fixation after 900 generations. But the pattern of results in the Battle of the Sexes and Leader games was quite different, with fixation of genes for defection following *STS* and cooperation following *TST*, indicating the evolution of coordinated alternating reciprocity.

The evidence for coordinated alternating reciprocity in the Battle of the Sexes and Leader is compelling. Nothing like this has been reported in evolutionary simulations of the Prisoner's Dilemma, and it is not difficult to see why this should be so. In the repeated Prisoner's Dilemma game, efficient reciprocity involves straightforward cooperation in response to cooperation and defection in response to defection, and the TFT strategy implements this efficiently. But in the repeated Battle of the Sexes and Leader games, reciprocity presents a more complicated and subtle challenge, and (TFT, TFT) is not in equilibrium. In these games, in order to improve on the expected payoffs of (2, 2) from mutual cooperation, the players need to evolve a form of reciprocity involving coordinated alternation between *C* and *D* in order to capitalize on the asymmetric (*C*, *D*) and (*D*, *C*) equilibrium points.

Genes encoding for defection following *STS* and cooperation following *TST* implement this type of alternating reciprocity efficiently. The *STS* three-move history implies that the player first cooperated while the co-player defected, then defected while the co-player cooperated, and then cooperated once again while the co-player defected, yielding a payoff stream of (4, 3), (3, 4), (4, 3), ... in Battle of the Sexes, or of (3, 4), (4, 3), (3, 4), ... in Leader. These payoffs are better for both players than those resulting from joint cooperation, namely (2, 2), (2, 2), (2, 2), ..., and in order to continue the more profitable payoff stream, Player I must defect on the next round. The same argument applies in reverse to the *TST* history. In this case, in order to continue the profitable payoff stream, Player I must cooperate on the following round.

Asymmetric alternation requires phase coordination to lock into the higher payoff streams in Battle of the Sexes or Leader. In human and animal social interaction, this is achieved by language and communication, facilitating a variety of coordination and turn-taking conventions. The data reported in this article appear to show, somewhat surprisingly, that it can evolve without the benefit of language or any other form of communication and without any insight on the part of the players into the strategic properties of the interaction. It is not obvious how this was achieved, but it is possible to envisage a plausible mechanism that could implement coordinated alternating reciprocity among players who lack insight and cannot communicate with each other.

The simplest implementation mechanism, similar to the 'CAD' mechanism suggested by Crowley (2001), is as follows. Consider first the situation that arises after an asymmetric (*C*, *D*) or (*D*, *C*) outcome. If every member of the population plays TFT, then, after an asymmetric outcome, the two players lock into coordinated alternating reciprocity and continue the alternating outcome sequence (*C*, *D*), (*D*, *C*), (*C*, *D*), (*D*, *C*), ... indefinitely. However, after a symmetric (*C*, *C*) or (*D*, *D*) outcome, TFT players repeat that symmetric outcome indefinitely. In neither case do they ever tap into the alternating outcome sequence, which is more profitable in certain games. If mixed strategies were possible, then the (*C*, *C*) lock-in could be avoided if, after every symmetric outcome, the players chose *C* with probability 1/2 and *D* with probability 1/2. In that case, after a symmetric (*C*, *C*) or (*D*, *D*) outcome, the probability of an asymmetric (*C*, *D*) or (*D*, *C*) outcome on the next round would be 1/2, and the pair would immediately enter the alternating outcome sequence. The probability of a symmetric (*C*, *C*) or (*D*, *D*) outcome on the following round would also be 1/2. After a symmetric outcome on round *t*, the players would lock into coordinated

alternating reciprocity with probability $1/2$ on round $t + 1$, with probability $3/4$ on round $t + 2$, and with probability $7/8$ on round $t + 2$. The mixed strategy that chooses C and D with equal probability maximizes these probabilities – any bias toward C or D reduces the probabilities of successful coordination.

Mixed strategies are not possible in a model such as the one described in this article, but the stochastic element required for this solution could arise at the population level. The population could evolve so that, after an asymmetric (C, D) or (D, C) outcome, all members of the population play TFT, but after a symmetric (C, C) or (D, D) outcome, 50 per cent of the population play C and 50 per cent play D . If these genes were distributed across the genomes independently for all loci associated with attainable three-move histories terminating in either (C, C) or (D, D), then almost 90 per cent of pairs selected from the population at random would lock into coordinated alternating reciprocity within three rounds.

The analysis above shows that coordinated alternating reciprocity cannot evolve in a uniform population, but it can evolve in a population in which genomes uniformly code for TFT after asymmetric outcomes and are maximally diverse after symmetric outcomes. If the payoff structure of the game favours coordinated alternating reciprocity, then the simplest population capable of evolving such interactive behaviour appears to be one whose members all play like TFT after asymmetric outcomes, half cooperating and half defecting after each attainable three-move history terminating in a symmetric outcome. This is probably how the coordinated alternating reciprocity evolved in the Battle of the Sexes and Leader simulations reported above. During the last 100 generations, the relative frequencies of genes coding for C choices were indeed close to $1/2$ for all three-move histories terminating in S or T (xyS or xyT).

Evolution was more stable in Battle of the Sexes and Leader than in the other games (Fig. 2). The explanation for this may lie in the strategic properties of the games and, more specifically, the degree to which the basic equilibrium points of the stage games satisfy the players' motivations. In Battle of the Sexes and Leader, there is no non-equilibrium outcome that either player prefers to the two equilibrium points. In the Prisoner's Dilemma game, on the other hand, each player prefers the out-of-equilibrium (C, C) outcome and one of the asymmetric outcomes to the (D, D) equilibrium point, and in Chicken, each player prefers the out-of-equilibrium (C, C) outcome to one of the equilibrium points. This may explain the greater evolutionary stability of Battle of the Sexes and Leader.

It is worth commenting, finally, on the fact that the Chicken simulation resulted in the evolution of symmetric reciprocity in spite of the fact that, like Battle of the Sexes and Leader, the basic Chicken stage game has only asymmetric equilibrium points in pure strategies. Chicken has equilibrium points at (C, D) and (D, C), but evolution tended to favour TFT-like strategies that cooperated most of the time. This was probably because joint cooperation is one of the additional equilibrium points that emerge in the repeated supergame, and with the payoff values that were used, it yields the same long-run expected payoff as coordinated alternation between (C, D) and (D, C). Considering the difficulty of evolving coordinated strategic alternation without the benefit of language, the evolution of strategies tending to generate (C, C) outcomes seems more likely a priori. The Chicken game is defined by the inequalities $T > R > S > P$ (Fig. 1E). Evolution of cooperation through coordinated alternating reciprocity might occur in a version of Chicken with $2R < S + T$ rather than $2R = S + T$, as in the version that was used. This is a problem for further research. The definitions of Battle of the Sexes ($S > T > R > P$) and Leader ($T > S > R > P$) imply $2R < S + T$, and coordinated alternating reciprocity is therefore always more profitable in the corresponding repeated supergames in these cases.

Acknowledgements

The authors wish to thank Alan Robinson and Robert Browning for help with programming and Ian Pountney for comments on an earlier draft of the article, the preparation of which was facilitated in part by a period of study leave granted to Andrew M. Colman by the University of Leicester and a research grant RES-000-23-0154 awarded by the Economic and Social Research Council.

References

- Axelrod, R. (1987). The evolution of strategies in the iterated Prisoner's Dilemma. In: *Algorithms and Simulated Annealing* (Davis, L., ed.) pp. 32-41. Los Alamos, CA: Morgan Kaufman.
- Axelrod, R., Dion, D. (1988). The further evolution of cooperation. *Science* 242, 1385-1390.
- Binmore, K. (1992). *Fun and Games: A Text on Game Theory*. Lexington, MA: Heath.
- Binmore, K., Samuelson, L. (1992). Evolutionary stability in repeated games played by finite automata. *J. Econ. Theor.* 57, 278-305.
- Bendor, J., Kramer, R., Stout, S. (1991). When in doubt: Cooperation in a noisy Prisoner's Dilemma. *J. Conflict Resolut.* 35, 691-719.
- Boerlijst, M. C., Nowak, M. A., Sigmund, K. (1997). The logic of contrition. *J. Theor. Biol.* 185, 281-293.
- Colman, A. M. (1995). *Game Theory and its Applications in the Social and Biological Sciences* (2nd edn). Oxford: Butterworth-Heinemann.
- Crowley, P. H. (2001). Dangerous games and the emergence of social structure: Evolving memory-based strategies for the generalized hawk-dove game. *Behav. Ecol.* 12, 753-760.
- Friedman, D. (1996). Equilibrium in evolutionary games: Some experimental results. *Econ. J.* 106, 1-25.
- Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*. Reading, MA: Addison-Wesley.
- Hofbauer, J., Sigmund, K. (1998). *Evolutionary Games and Population Dynamics*. Cambridge, UK: Cambridge University Press.
- Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems* (2nd edn). Ann Arbor, MI: University of Michigan Press.
- Kraines, D., Kraines, V. (1989). Pavlov and the Prisoner's Dilemma. *Theor. Decis.* 26, 47-79.
- Kraines, D., Kraines, V. (1995). Evolution of learning among Pavlov strategies in a competitive environment with noise. *J. Conflict Resolut.* 39, 439-466.
- Leimar, O. (1997). Repeated games: A state space approach. *J. Theor. Biol.* 184, 471-498.
- Lindgren, K. (1991). Evolutionary phenomena in simple dynamics. In: *Artificial Life II: Studies in Complexity* (Langton, C. G. Taylor, C., Farmer, J. D., Rasmussen, S., Eds) vol. 10, pp. 295-312, Reading, MA: Addison-Wesley.
- Lipman, B. L. (1986). Cooperation among egoists in Prisoner's Dilemma and Chicken games. *Public Choice* 51, 315-331.
- Lorberbaum, J. (1994). No strategy is evolutionarily stable in the repeated Prisoner's Dilemma. *J. Theor. Biol.* 168, 117-130.
- Luce, R. D., Raiffa, H. (1957). *Games and Decisions: Introduction and Critical Survey*. New York: Wiley.
- Maynard Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge, UK: Cambridge University Press.
- Maynard Smith, J., Price, G. R. (1973). The logic of animal conflict. *Nature* 246, 15-18.
- Nowak, M. A., May, R. M. (1992). Evolutionary games and spatial chaos. *Nature* 359, 826-829.
- Nowak, M. A., May, R. M., Sigmund, K. (1995). The arithmetics of mutual help. *Sci. Am.* 272(6), 50-55.
- Nowak, M. A., Sigmund, K. (1992). Tit for tat in heterogeneous population. *Nature* 355, 250-253.
- Nowak, M. A., Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature* 364, 56-58.
- Nowak, M. A., Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature* 393, 573-577.
- Posch, M. (1999). Win-stay, lose-shift strategies for repeated games: Memory length, aspiration levels and noise. *J. Theor. Biol.* 198, 183-195.
- Rapoport, A. (1967). Exploiter, Leader, Hero, and Martyr: The four archetypes of the 2×2 game. *Behav. Sci.* 12, 81-84.
- Rapoport, A., Guyer, M. (1966). A taxonomy of 2×2 games. *Gen. Systems* 11, 203-214.
- Riolo, R. (1992). Survival of the fittest bits. *Sci. Am.* 267(7), 114-116.
- Roth, A., Erev, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games Econ. Behav.* 8, 164-212.
- Russell, B. (1959). *Common Sense and Nuclear Warfare*. New York: Simon Schuster.
- Samuelson, L. (1998). *Evolutionary Games and Equilibrium Selection*. Cambridge, MA: MIT Press.

Wu, J., Axelrod, R. (1995). How to cope with noise in the iterated Prisoner's Dilemma. *J. Conflict Resolut.* 39, 183-189.