Evolutionary Ecology Research, 2009, **11**: 949–963

Evolution of cooperative turn-taking

Andrew M. Colman and Lindsay Browning

School of Psychology, University of Leicester, Leicester, UK

ABSTRACT

Question: How can the evolution of turn-taking be explained in species without language?

Features of model: Using a genetic algorithm incorporating mutation and crossover, we studied noisy decision making in three repeated two-player games in which we predicted on theoretical grounds that cooperative turn-taking would evolve and three games in which we expected synchronized cooperation to evolve.

Ranges of key variables: We set population size to 20, number of rounds to be played by each pair in each generation to 200, and number of evolutionary generations to 2000, and we repeated each simulation 10 times to check the stability of the results.

Results: Cooperative turn-taking and (unexpectedly) a form of double turn-taking evolved in the alternation games, and joint cooperation evolved in the synchronization games. We propose a mechanism to explain how cooperative turn-taking can evolve mechanically, even without communication or insight, as it did in our simulations.

Keywords: Battle of the Sexes, cooperation, coordination game, evolutionary game, genetic algorithm, Prisoner's Dilemma, Stag Hunt, tit for tat, turn-taking.

INTRODUCTION

It has been suggested that 'the most important unanswered question in evolutionary biology, and more generally in the social sciences, is how co-operative behaviour evolved and can be maintained in human and other animal groups and societies' (May, 2006, p. 109). The introduction of genetic algorithms (Holland, 1975; Goldberg, 1989) provided new techniques for investigating the natural selection of cooperation in models far too complicated for analytical solution. Studies of the evolution of cooperation using these techniques have focused almost exclusively on synchronized cooperation in the repeated (or iterated) Prisoner's Dilemma game (e.g. Axelrod, 1987; Axelrod and Dion, 1988; Kraines and Kraines, 1989, 1995; Bendor *et al.*, 1991; Lindgren, 1991; Nowak and May, 1992; Nowak and Sigmund, 1992, 1993, 1998; Fogel, 1993; Lorberbaum, 1994; Nowak *et al.*, 1995; Wu and Axelrod, 1995; Worden and Levin, 2007) or closely related games (Lipman, 1986; Binmore and Samuelson, 1992; Roth and Erev, 1995; Friedman, 1996; Posch, 1999; Crowley, 2001).

Correspondence: A.M. Colman, School of Psychology, University of Leicester, Leicester LE1 7RH, UK. e-mail: amc@le.ac.uk

Consult the copyright statement on the inside front cover for non-commercial copying policies.

© 2009 Andrew M. Colman

$$\begin{array}{cc|c|c|}
 & & \multicolumn{2}{c}{\text{II}} \\
 & & C & D \\
\cline{3-4}
\text{I} & C & 3,\ 3 & 0,\ 5 \\
\cline{3-4}
 & D & 5,\ 0 & 1,\ 1 \\
\cline{3-4}
\end{array}
\qquad
\begin{array}{cc|c|c|}
 & & \multicolumn{2}{c}{\text{II}} \\
 & & C & D \\
\cline{3-4}
\text{I} & C & R, R & S, T \\
\cline{3-4}
 & D & T, S & P, P \\
\cline{3-4}
\end{array}$$

Fig. 1. Prisoner's Dilemma game and generic payoff matrix. Left panel: the standard Prisoner's Dilemma game. Right panel: a generic payoff matrix for any symmetric 2 × 2 game.

The Prisoner's Dilemma (Fig. 1, left panel) is a natural game for studying synchronized cooperation, because it models strategic interactions in which each player has a choice between two actions, usually labelled cooperation and defection (*C* and *D*) respectively, with unilateral defection yielding the best payoff to the defector and the worst to the cooperator, but with both players better off if they cooperate than if both defect. Player I chooses row *C* or *D*, Player II independently chooses column *C* or *D*, and the two numbers in each cell are the payoffs to Players I and II respectively for that outcome. In the one-shot (unrepeated) version of the game, cooperation is clearly irrational, because 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, hence a player receives a higher payoff by defecting than by cooperating whatever the co-player does. Furthermore, the (*D*, *D*) outcome is the only *Nash equilibrium* of the one-shot Prisoner's Dilemma game, a Nash equilibrium being a combination of strategies that are *best replies* to each other, in the sense that each yields the highest payoff to the player choosing it, given the strategy chosen by the co-player. In view of the dominance of the *D* strategies and the fact that they combine to form a unique Nash equilibrium, it is generally accepted that defection is the only rational strategy in the one-shot Prisoner's Dilemma game. Nevertheless, paradoxically, both players receive higher payoffs if both cooperate than if both defect.

Although investigations based on the Prisoner's Dilemma have yielded important findings, many questions remain unanswered. In particular, the evolution of cooperative turn-taking poses a special problem of explanation, because the Prisoner's Dilemma models only synchronized cooperation and is unsuited to the analysis of turn-taking. Coordinated turn-taking is clearly a form of cooperative behaviour, but it remains unexplained by research based on the Prisoner's Dilemma game and cannot even be modelled satisfactorily within the strategic framework of that game. However, there are many naturally occurring interactions in which individuals can maximize their payoffs only by coordinated turn-taking. For example, dyadic allogrooming behaviour in non-human primates (Manson *et al.*, 2004) and in many other species, including neonatal antelopes (Mooring and Hart, 1997), typically involves coordinated turn-taking. In several species of Antarctic penguins, mating pairs take turns foraging at sea while the other member of the pair incubates eggs (Weimerskirch *et al.*, 1992) or remains with newly hatched chicks (Trivelpiece *et al.*, 1987). Human conversations necessarily entail turn-taking (O'Connell *et al.*, 1990), and conversational turn-taking is known to be controlled by at least six distinct cues (Duncan, 1972). Similar turn-taking has been observed in non-verbal interaction, even between 6-week-old infants and adults (Crown *et al.*, 1993). In all of these cases, and many others, turn-taking confers a clear advantage to individuals, either directly or in terms of Darwinian fitness. Nonetheless, although reciprocity has long been recognized as a significant aspect of cooperation (e.g. Trivers, 1971; Axelrod and Hamilton, 1981; Nowak and Sigmund, 1994), published research on the evolution of cooperation has been restricted almost exclusively to synchronized cooperation in games that provide no payoff advantage for turn-taking.

The distinctive strategic properties of the Prisoner's Dilemma game have imposed an artificial limitation on research into the evolution of cooperation. One property in particular renders it unsuitable for modelling turn-taking. Figure 1 (right panel) shows a generic payoff matrix for symmetric 2×2 games. The standard Prisoner's Dilemma game (left panel) is defined by the inequalities $T > R > P > S$, and the additional condition $2R > S + T$ that many game theorists consider to be part of its definition (e.g. Rapoport and Chammah, 1965; Axelrod and Hamilton, 1981; Hofstadter, 1983; Axelrod, 1984; Fogel, 1993; Frean, 1994). In the repeated Prisoner's Dilemma, this inequality implies that players cannot gain any payoff advantage by taking turns, and coordinated turn-taking behaviour cannot be interpreted or even meaningfully defined within its strategic framework. It ensures that the payoff stream $S + T + S + \ldots$ that would result from coordinated turn-taking does not exceed the payoff stream resulting from synchronized cooperation, namely $R + R + R + \ldots$. This applies even to so-called *asymmetric* Prisoner's Dilemmas (e.g. Sheposh and Gallo, 1973; Croson, 1999), and it applies even to the *alternating* Prisoner's Dilemma (Nowak and Sigmund, 1994), in which the payoff structure is as shown in Fig. 1 (left panel) but players alternate in choosing C or D rather than choosing simultaneously. Given these payoffs, coordinated turn-taking generates the payoff stream $5 + 0 + 5 + 0 + \ldots$, with a mean of 2.50, whereas synchronized cooperation yields $3 + 3 + 3 + \ldots$, with a mean of 3.00. The Prisoner's Dilemma game provides a useful model for many strategic interactions, but the condition $2R > S + T$ precludes an advantage for coordinated turn-taking.

Apart from Nowak and Sigmund (1994), a small number of more recent studies have addressed the problem of turn-taking. In particular, Neill (2003) investigated a version of the repeated Prisoner's Dilemma game with $2R < S + T$ and, using specialized techniques, found several evolutionarily successful strategies for playing it. Browning and Colman (2004) carried out a preliminary investigation of several games with asymmetric Nash equilibria and found some indirect evidence for the evolution of coordinated alternation. However, they did not study the prototypic Battle of the Sexes game, with zero payoffs in the non-equilibrium cells, although this game affords the simplest and clearest test opportunity for turn-taking. Furthermore, their simulations did not incorporate noise, an important factor that generally facilitates evolutionary processes (Blume, 2003). Above all, their research methods did not enable actual turn-taking choice sequences to be observed and counted, hence their evidence, based on an examination of properties of the evolving genotypes over generations, was only indirect and circumstantial.

In this article, we report the results of a study using a genetic algorithm incorporating mutation, crossover, and noise to investigate the evolution of cooperation in the simplest games in which players can maximize their payoffs only by coordinated turn-taking, including the prototypic Battle of the Sexes game. The results, based on direct counts of turn-taking sequences, show the evolution of cooperative turn-taking and double turn-taking, but only in games in which it is profitable to the players. We suggest a new mechanism capable of explaining how coordinated turn-taking can evolve without communication or insight and without the use of randomized mixed strategies.

EXPERIMENTAL AND CONTROL GAMES

We focused on three symmetric 2×2 coordination games with $R \geq P$ and $2R < S + T$ (see Fig. 1, right panel). When such games are repeated, cooperation may be defined in terms of coordinated turn-taking, because that is the only way in which both players maximize their payoffs, given that the payoff stream $(R + R + R + \dots) < (S + T + S + \dots)$. Each of the games that we used has asymmetric Nash equilibria at (C, D) and (D, C), and in each game one player prefers (C, D) to (D, C) whereas the other prefers (D, C) to (C, D). These are thus coordination games, and coordinated turn-taking is evidently a form of cooperation, because both individuals must play their parts to achieve payoff maximization through alternating S and T payoffs. We call these *alternation games*, and we hypothesized that coordinated turn-taking would evolve when they were repeated many times in an appropriate genetic algorithm. As benchmarks for comparison, we also studied three control games with $R \geq P$ and $2R > S + T$. We call these *synchronization games*, because synchronized cooperation achieves payoff maximization when repeated, and we hypothesized that such behaviour would evolve in these games.

Alternation games

The experimental (alternation) games are shown in Fig. 2 (top row). The Battle of the Sexes game (defined by the relations $T > S > R = P$) was introduced by Luce and Raiffa (1957, pp. 90–94). It has Nash equilibria at (C, D) and (D, C), with Player I preferring (D, C) and Player II (C, D). The stripped-down version shown in Fig. 2, with zero payoffs in the non-equilibrium cells, is the simplest possible alternation game, without any distracting features. The Hero game $(S > T > R > P)$ and Leader game $(T > S > R > P)$ were first explicitly identified and named by Rapoport (1967), and both have Nash equilibria at (C, D) and (D, C), but in Hero Player I prefers (C, D) and Player II (D, C), whereas in Leader Player I prefers (D, C) and Player II (C, D). Hero is a game in which each player receives the highest possible payoff when the *other* player is the sole defector, accepting the second-best payoff and in this sense playing the role of hero, whereas Leader is a pre-emption game in which

		II	
		C	D
I	C	0, 0	1, 2
	D	2, 1	0, 0

Battle of the Sexes

		II	
		C	D
I	C	2, 2	4, 3
	D	3, 4	1, 1

Hero

		II	
		C	D
I	C	2, 2	3, 4
	D	4, 3	1, 1

Leader

		II	
		C	D
I	C	3, 3	0, 5
	D	5, 0	1, 1

Prisoner's Dilemma

		II	
		C	D
I	C	3, 3	0, 2
	D	2, 0	2, 2

Rousseau's Stag Hunt

		II	
		C	D
I	C	9, 9	0, 8
	D	8, 0	7, 7

Aumann's Stag Hunt

Fig. 2. Games with and without turn-taking opportunities, arranged with $R \geq P$. The top row shows alternation games with $2R < S + T$. The bottom row shows synchronization games with $2R > S + T$.

each player prefers to be the sole defector while the other player accepts the second-best payoff. In the indefinitely repeated versions of these three alternation games, the payoff stream $S + T + S + \ldots$ that a player receives under coordinated turn-taking is preferable to the payoff stream $R + R + R + \ldots$ from synchronized cooperation.

Synchronization games

The synchronization games that we used for comparison, shown in Fig. 2 (bottom row), were the Prisoner's Dilemma game ($T > R > P > S$), a game with a unique Nash equilibrium at (D, D), and two sharply distinct versions of the Stag Hunt game. All three synchronization games have $2R > S + T$. The Stag Hunt games are named after Rousseau's (1755, Part II, paragraph 9) interpretation, in which two hunters have to cooperate to catch a stag, but each is tempted to chase after the lesser prize of a hare, which can be caught without the other's cooperation. Rousseau's original Stag Hunt game, introduced into the literature of game theory by Lewis (1969, p. 7), has $R > T = P > S$ and Nash equilibria at (C, C) and (D, D), both players preferring (C, C) to (D, D). Aumann's Stag Hunt game ($R > T > P > S$), introduced by Aumann (1987) and analysed in depth by Harsanyi and Selten (1988, pp. 357–359), is a sharper version, again with Nash equilibria at (C, C) and D, D), but with greater relative risk and less relative advantage associated with the C strategy. In both Stag Hunt games, the (C, C) equilibrium is *payoff-dominant*, because it yields higher payoffs to both players than the (D, D) equilibrium, but the (D, D) equilibrium is *risk-dominant*, because it avoids the risk of the worst possible payoff (zero). More precisely, using the notation in Fig. 1 (right panel), the equilibrium (C, C) risk-dominates the equilibrium (D, D) if $(R - T)^2 > (P - S)^2$, and the equilibrium (D, D) risk-dominates (C, C) if the reverse inequality holds. In indefinitely repeated versions of all three synchronization games, $(R + R + R + \ldots) > (S + T + S + \ldots)$, and consequently players gain no payoff advantage from taking turns.

GENETIC ALGORITHM

Each of our experimental simulations began with a population of 20 random programmed strategies for playing repeated 2×2 games. A programmed strategy is simply a finite state automaton or Moore machine (Moore, 1957), generating on each round (each repetition of the game) a unique move (C or D) for every possible history of three immediately preceding rounds. Because an individual payoff on a single round is P, R, S or T, depending on both players' choices, there are $4^3 = 64$ distinguishable three-round histories. The programmed strategies in our study were accordingly encoded in 64-bit strings. Every possible programmed strategy with a memory restricted to three moves can be represented in this way, and there are $2^{64} \approx 10^{19}$ such programmed strategies. We used 0 to denote a D move and 1 a C move. For example, a programmed strategy with the binary digit 1 associated with the three-round history RRR responds to three consecutive R payoffs – three consecutive (C, C) outcomes – by playing C, and if it has the digit 0 in its 64-bit string at the locus associated with the three-round history RRP, then it plays D after two joint cooperative outcomes followed by a joint defection, and so on. Because the three-round histories are recorded as strings of payoffs encoded as P, R, S or T on each round, no information about strategy choices is lost, because a payoff tagged in this way corresponds to a unique strategy pair chosen by the players on the corresponding round, even if there is another payoff in the game with the same numerical value. A programmed strategy was represented by a string of

64 zeros and ones, each digit encoding a strategy choice for a unique three-round history corresponding to its position or locus in the string. To get each game started before any three-round history was well defined, we added six arbitrary binary digits, called *premise genes*, to create an arbitrary and variable initial three-round history, making a total 70-bit string for a complete specification of an individual's genome.

In each evolutionary generation, every member of the population of 20 programmed strategies was paired with every other member to play 200 rounds of the game. A genetic algorithm then formed offspring programmed strategies for the next generation from the most successful members of the current generation. Our specially designed genetic algorithm, based on Holland (1975), was implemented in the following 10 steps:

1. Assign payoff values according to one of the games in Fig. 2.
2. Set population size to 20, number of rounds to be played by each pair in each generation to 200, and number of evolutionary generations to 2000.
3. Generate a random 70-digit binary number for each individual in the population.
4. In every generation, pair each individual with each of the others to play 200 rounds, assign payoffs according to the game's payoffs, and sum payoffs for each individual over the 200 rounds for all pairings.
5. Introduce random misimplementation noise: for every C or D move generated from a three-round history, substitute the alternative move with probability $p = 0.001$.
6. Compute each individual's total payoff over its 19 pairwise interactions and assign a mating probability proportional to its total payoff.
7. For each offspring, select two parents randomly in proportion to mating probability.
8. Simulate crossover by randomly selecting a number from 3 through 67, breaking each parental genome at that point, and combining the code before the crossover point from one parent and after the crossover point from the other to form the offspring.
9. Simulate mutation by altering each digit of the offspring genome with probability $q = 0.01$ at the start of each generation.
10. Replace 20 parents with 20 offspring at each new generation, and loop back to repeat steps 4 to 9 for 2000 generations.

We replicated each simulation 10 times to check the stability of the results, and our simulations all turned out to be reasonably stable across replications, although the Prisoner's Dilemma simulation was slightly more variable across replications than the others, presumably because it is highly unusual in having no evolutionarily stable strategy (Lorberbaum, 1994).

RESULTS

The most striking feature of the results is the comparative instability of the Prisoner's Dilemma simulation relative to the other five simulations. Choice behaviour in the three alternation games all evolved rapidly to mean payoffs per round close to the predicted optimum of $(S + T)/2$ expected under coordinated turn-taking (1.50 in Battle of the Sexes and 3.50 in Hero and Leader simulations), and remained close to the optimum for the remaining generations. In the synchronization games, behaviour also evolved close to the predicted optimum of R in all games (3.00 in the Prisoner's Dilemma game and Rousseau's Stag Hunt game and 9.00 in Aumann's Stag Hunt game), although large deviations

Table 1. Percentages of distinct outcomes and reciprocity indices in generation 2000 of each simulation, averaged across replications

Game	(C, C)	(C, D)	(D, C)	(D, D)	Recip.
Battle of the Sexes	4.59	44.78	44.78	5.85	80.59
Hero	9.79	42.69	42.69	4.83	84.98
Leader	6.97	42.67	42.67	7.69	66.70
Prisoner's Dilemma	48.49	9.56	9.56	32.39	52.42
Rousseau's Stag Hunt	96.97	1.20	1.20	0.63	97.39
Aumann's Stag Hunt	88.10	1.59	1.59	8.72	90.17

Note: Percentages of (C, D) and (D, C) are identical because a (C, D) outcome for one player always corresponds to a (D, C) outcome for the other.

to suboptimal payoffs occurred frequently in the Prisoner's Dilemma game, no doubt reflecting the dominance of the D strategy and the (D, D) Nash equilibrium. These findings were typical of the 10 replications of each simulation, and they are fully consistent with our hypotheses.

Table 1 shows the percentages of joint cooperative (C, C), asymmetric (C, D) and (D, C), and joint defecting (D, D) outcomes in generation 2000 of each simulation, averaged across replications in each case. The right-hand column shows the reciprocity indices in generation 2000, also averaged across replications, defined as the percentage of rounds on which players repeated the strategy (C or D) chosen by the co-player on the previous round. The reciprocity index is high not only when players reciprocate C or D choices directly, as in sequences such as (C, C), (C, C), (C, C), . . . , but also in turn-taking sequences such as (C, D), (D, C), (C, D), . . . , in which each player also reciprocates the co-player's strategy from the previous round – a form of reciprocity that has generally been overlooked and cannot even be conceptualized within the framework of a synchronization game. These data show high levels of reciprocity in all games except the Prisoner's Dilemma game, in which (C, C) and (D, D) outcomes predominated but were immediately reciprocated in only 52% of rounds. Behaviour was much more unstable in the Prisoner's Dilemma than in the other simulations, presumably because of the paradoxical strategic structure of the game and the difficulty in defining an optimal programmed strategy for the indefinitely repeated version in the absence of any evolutionarily stable strategy. In the other five games, strategy choices were immediately reciprocated on more than two-thirds of rounds in generation 2000. In the three alternation games, this reciprocation appears from the columns headed (C, D) and (D, C) to have taken the form of sequences such as (C, D), (D, C), (C, D), . . . , because most outcomes were either (C, D) or (D, C). Reciprocity was slightly lower in the Leader than the other turn-taking simulations, and this will be discussed below. In the synchronization games, on the other hand, the reciprocation appears to have involved sequences such as (C, C), (C, C), (C, C), . . . , and in the Prisoner's Dilemma game also (D, D), (D, D), (D, D), ... These data are also fully consistent with our hypotheses.

To provide direct evidence of turn-taking, we focus in Table 2 on counts of three-round histories in selected generations. These three-round histories are actual outcomes, representing individual choice pairs, counted in the selected generations, and they provide unambiguous evidence for the evolution of turn-taking behaviour. Using x and y as variables representing arbitrary values P, R, S, and T, all three-round histories culminating

Table 2. Prevalence of three-round histories (percentages), averaged across replications of each simulation in generations 1, 100, 1000, and 2000

Generation	*xyP*	*xyR*	*xyS*	*xyT*
Battle of the Sexes				
1	24.41	24.38	25.62	25.60
100	8.36	5.20	43.12	43.32
1000	8.51	6.74	42.37	42.37
2000	5.88	4.67	44.78	44.66
Hero				
1	24.76	24.57	25.33	25.34
100	4.43	10.86	42.48	42.23
1000	4.94	6.70	44.10	44.26
2000	4.96	9.80	42.71	42.52
Leader				
1	24.22	25.27	25.26	25.25
100	11.91	8.61	39.69	39.79
1000	7.67	9.68	41.28	41.36
2000	7.80	7.05	42.60	42.55
Prisoner's Dilemma				
1	23.40	25.21	25.71	25.68
100	24.67	53.63	10.80	10.91
1000	23.62	63.18	6.55	6.64
2000	32.33	48.44	9.61	9.62
Rousseau's Stag Hunt				
1	25.50	23.74	25.38	25.38
100	17.72	78.57	1.82	1.89
1000	1.75	96.01	1.17	1.08
2000	0.63	96.77	1.34	1.26
Aumann's Stag Hunt				
1	23.30	27.18	24.74	24.78
100	19.11	78.37	1.25	1.26
1000	3.97	93.07	1.46	1.50
2000	8.73	87.94	1.65	1.68

Note: The symbols x and y are variables representing arbitrary values P, R, S or T from Fig. 1 (right panel).

in P are labelled xyP, and similarly for R, S, and T. A payoff of P results only when both players choose D on the same round (see Fig. 1, right panel); therefore, any xyP triplet indicates synchronized D moves on that round, and for an analogous reason xyR indicates synchronized C moves. In this sense, P is the *complementary* payoff of P, and R is the complementary payoff of R. On the other hand, xyS and xyT indicate asymmetric

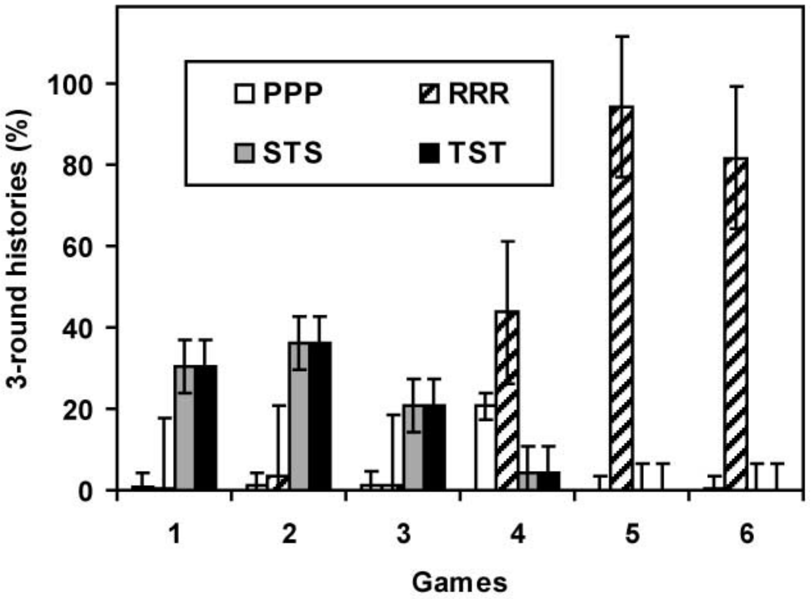

Fig. 3. Mean frequencies of three-round histories and in generation 2000 in simulations of six games. Error bars indicate standard errors of the means.

behaviour, one player choosing *C* while the other chooses *D*, hence the complementary payoffs of *S* and *T* are *T* and *S* respectively – whenever one player receives a payoff of *S*, the co-player receives *T*, and vice versa.

The results shown in Table 2 are clearly consistent with rapid evolution of asymmetric strategies in the Battle of the Sexes, Hero, and Leader simulations, and the evolution of symmetric strategies – mainly synchronized *C* choices – in the Prisoner's Dilemma and both Stag Hunt simulations. After 2000 generations of evolution, the total (summed) percentages of asymmetric *xyS* or *xyT* triplets were 89.44% in the Battle of the Sexes simulations, 85.23% in the Hero simulations, and 85.15% in the Leader simulations – the alternation games – compared with 19.23% in the Prisoner's Dilemma simulations, and only 2.60% and 3.33% respectively in Rousseau's and Aumann's Stag Hunt simulations. As we anticipated, evolution in the synchronization games was predominantly towards symmetric *C* choices, although less consistently so in the Prisoner's Dilemma game.

The relative frequencies of specific three-round histories played in generation 2000 provide the first direct evidence of cooperative turn-taking (see Fig. 3). A count of all 64 possible three-round histories revealed that, in the Battle of the Sexes game – the prototypic alternation game – across the 10 simulations, the most prevalent triplets by far were *TST* (30.57%) and *STS* (30.56%), together accounting for 61.13% of all three-round histories in the final generation. In view of the payoff complementarity of *T* and *S*, a *T* or *S* payoff occurring only when one player chooses *D* while the other chooses *C* (see Fig. 1, right panel), these triplets clearly reflect coordinated turn-taking, because one player's *T* always co-occurs with the co-player's *S*, and the percentages of *TST* and *STS* are always the same, within the limits of random mutation in the genetic algorithm's premise genes, which fluctuated without complementary outcome pairing.

In the Battle of the Sexes simulations, the only other three-round histories that evolved to significant (≥ 5%) prevalence in generation 2000 were *TTS* (5.64%), *SST* (5.61%), *TSS* (5.49%), and *STT* (5.44%), accounting for another 22.18% of the three-round histories in the final generation. These triplets indicate coordinated *double* turn-taking, because both *TTS* and *SST* imply that one player chose *D* twice before switching, while the other chose *C* twice before switching. We did not anticipate these triplets, but they clearly represent a familiar form of coordinated turn-taking – 'I'll walk the dog today and tomorrow if you do it on the following two days, then I'll do it on the two days after that, and so on' – and their evolution is remarkable. The Hero and Leader games yielded similar results. In the Hero simulations, the most prevalent triplets were *TST* (36.18%) and *STS* (36.16%), indicating coordinated turn-taking, and no other triplets evolved to significant prevalence. In the Leader simulations, the most prevalent triplets were *STS* (20.87%) and *TST* (20.84%), indicating coordinated turn-taking, followed by *TTS* (7.68%), *SST* (7.68%), *TSS* (7.36%), and *STT* (7.36%), all indicating coordinated double turn-taking. The large proportion of double turn-taking in the Leader simulations may help to explain why the reciprocity index for this game in Table 1 was slightly lower than for other alternation games.

In the synchronization games, none of the turn-taking or double turn-taking triplets evolved to 5% or greater prevalence in generation 2000. In the Prisoner's Dilemma game, only *RRR* (43.77%), indicating synchronized *C* choices, and *PPP* (20.62%), indicating synchronized *D* choices, evolved to significant prevalence. In Rousseau's Stag Hunt game, only *RRR* (94.35%), and in Aumann's Stag Hunt game, also only *RRR* (81.78%), evolved to significant prevalence. In both Stag Hunt games, behaviour evolved to the payoff-dominant (*C*, *C*) equilibria, not the risk-dominant (*D*, *D*) equilibria predicted by some theories of evolutionary games (e.g. Young, 1998). The triplet counts for all six games corroborate our hypotheses strongly and directly.

DISCUSSION

The results reveal that cooperative turn-taking can evolve without communication between players, even in the austere social environment of Moore machines, and that it evolves only in games in which it is profitable to the players. Dynamic forms of cooperation, including turn-taking, are ubiquitous in human and animal interaction, and their evolution can be studied within a framework of appropriate coordination games and genetic algorithms. We defined alternation games as those in which $2R < S + T$ (see Fig. 1, right panel), with the consequence that players can benefit by taking turns, and we argued that in such games cooperation should be defined in terms of turn-taking. Turn-taking evolved prolifically in all three alternation games that we investigated. The simplest possible example is the prototypic Battle of the Sexes game, with zero payoffs for both players in non-equilibrium outcomes. In our 10 simulations of this game, turn-taking and double turn-taking, taken together, accounted for a remarkable 83.31% of three-round histories in the players' strategy choices occurring after 2000 generations of evolution.

Results for the other alternation games were almost equally impressive. Turn-taking and double turn-taking accounted for 72.34% of three-round histories in the Hero simulations and 71.79% in the Leader simulations in the final generation. In comparison, in the synchronization control games with $2R > S + T$, turn-taking did not evolve significantly, but joint cooperation and joint defection (mostly joint cooperation) evolved to 64.39% of three-round histories in the Prisoner's Dilemma simulations, 94.35% in Rousseau's Stag

Hunt simulations, and 81.78% in Aumann's Stag Hunt simulations in the final generation. It is also noteworthy that no other three-round history – none apart from turn-taking and double turn-taking in the alternation games and none apart from joint cooperation and joint defection in the synchronization games – was significantly prevalent after 2000 generations of evolution in any of the games that we studied.

Taken together, these results provide strong evidence in support of our hypothesis that coordinated turn-taking would evolve in games in which $2R < S + T$ and in those games only. This finding is gratifyingly clear-cut, but it presents a challenging problem of explanation and understanding. Turn-taking requires precise coordination and, in human strategic interaction, such behaviour is facilitated by a variety of linguistic and non-linguistic turn-taking conventions. However, it is clear from the data reported in this article that it can evolve without communication or insight on the part of the players, and without the use of randomization (mixed strategies). In our simulations, the players were Moore machines programmed merely to generate a unique move in response to any possible sequence of three preceding outcomes. How could such robotic and uncommunicative players evolve turn-taking propensities?

We begin by noting that the tit-for-tat (TFT) programmed strategy for indefinitely repeated Prisoner's Dilemma games (e.g. Axelrod, 1984, 1987) does not solve the problem. TFT is a simple strategy based on a one-move memory for the co-player's strategy, in contrast to our genetic algorithm, which is based on a three-move memory for previous outcomes of the game. TFT plays C on Round 1, then on every subsequent round t plays the strategy played by your co-player on round $t - 1$. In the Prisoner's Dilemma game, a pair of TFT players cooperate on every round. In the Battle of the Sexes (Fig. 2, top left), after an asymmetric (C, D) or (D, C) outcome, a pair of TFT players immediately lock into the mutually profitable (C, D), (D, C), (C, D), . . . turn-taking pattern with payoff streams 1, 2, 1, 2, . . . , repeated indefinitely. But after a symmetric (C, C) or (D, D) outcome, play oscillates between (C, C) and (D, D), with the unprofitable payoff stream 0, 0, 0, 0, ... If these players were able to use a uniform mixed strategy (choosing C or D randomly with equal probabilities) after every symmetric outcome, then the probability that they would enter the asymmetric turn-taking sequence, using TFT, would be 1/2 on Round $t + 1$, 3/4 on Round $t + 2$, 7/8 on Round $t + 3$, and so on. This mixed strategy $(1/2C, 1/2D)$ maximizes the probabilities and minimizes the waiting time for entering the turn-taking sequence and beginning to use TFT.

However, our players were constrained to choose pure strategies only; they could not and did not use mixed strategies, and there is evidence that even human players are unable to choose randomly without the help of randomizing devices (e.g. Rapoport and Budescu, 1997; Rosenthal *et al.*, 2003). The challenge is to explain how cooperation can evolve without randomization. Our suggestion is that the *individual players* in our simulations evolved to play TFT after asymmetric outcomes and that the *population* evolved to a genetic dimorphism in which 50% of the players invariably chose C and 50% invariably chose D after symmetric outcomes. Consider once again the Battle of the Sexes (Fig. 2, top left), and suppose that Player I chooses C on Round t. Player II plays C on Round t with probability 1/2, in which case the outcome is (C, C). After this symmetric outcome, each player is independently programmed to choose either C or D with certainty on the following round, depending on its type in the dimorphic population. The outcome on Round $t + 1$ is therefore one of four equiprobable outcomes, (C, C), (C, D), (D, C), or (D, D), and the conditional probability that the players choose different moves on Round $t + 1$ is 2/4 = 1/2. It follows that the joint probability of Player II choosing C on Round t and either of the outcomes (C, D) or (D, C) occurring on Round $t + 1$ is $(1/2)(1/2) = 1/4$. The other possibility on Round t is that Player II chooses D. In that case, the outcome on Round t is (C, D), and both players play TFT with certainty, so that the outcome on Round $t + 1$ is (D, C). Because Player II is equally likely to choose C or D on Round t, the total probability of a (C, D) or (D, C) outcome on Round $t + 1$, conditioned on Player I choosing C on Round t, is $1/4 + 1/2 = 3/4$ and, by symmetry, the same probability applies if Player I chooses D on Round t. This provides an explanation for how cooperation can evolve without mixed strategies, up to a population prevalence of 75%, in the repeated Battle of the Sexes, and it applies equally to other alternation games. This mechanism, or something functionally equivalent to it, seems necessary to explain the evolution of cooperative turn-taking.

A crucial test of this proposed mechanism is whether populations did indeed evolve to 50–50 dimorphisms having genes coding for C or D choices after xyR and xyP three-round histories associated with symmetric outcomes. In our simulations, the percentages of Moore machines in generation 2000 with genes coding for C after xyR are close to 50% in all three alternation games (48% in Battle of the Sexes, 53% in Hero, and 51% in Leader), and the percentages coding for C choices after xyP are also close to 50% (48% in Battle of the Sexes, 54% in Hero, and 50% in Leader). Within the limits of the high mutation rate built into the genetic algorithm, these data strongly support the theoretical predictions.

In conclusion, we comment on some indirectly relevant previous contributions, in chronological order. First, Aumann's (1974, 1987) concept of *correlated equilibrium* is interestingly related to coordinated turn-taking. Correlated equilibrium is a form of coordination for choosing between multiple Nash equilibria in a game by linking strategy choices to an external event, such as a coin toss. In the Battle of the Sexes game shown in Fig. 2, for example, human players might agree to choose the (C, D) equilibrium if a coin falls heads and (D, C) if it falls tails. If the coin is fair, then each player's expected payoff is $(1 + 2)/2 = 1.50$, and this is the average payoff that a player can expect, in the limit, if the game is repeated indefinitely. However, this is not turn-taking behaviour. It requires not only an external coordinating device but also a prior agreement between the players, and both are difficult to interpret in non-human turn-taking behaviour. Furthermore, it does not guarantee equal payoffs over several rounds, especially over a short run, because a fair coin does not usually yield exactly equal numbers of heads and tails, nor does it maximize an individual's aggregate payoff.

Second, Rapoport *et al.* (1976, Chapters 9, 11) reported the results of several experiments in which significant numbers of human dyads learned rapidly to alternate between strategy choices in 2×2 games with $2R < S + T$. This provides useful evidence that human turn-taking occurs not only in naturally occurring interactions, but also in abstract experimental games, although this has no direct bearing on evolution based on natural selection. In adaptive learning, individual players modify their strategy choices on the basis of experience; in evolutionary adaptation, strategy choices change only through random mutation, and natural selection increases the relative frequency in the population of the most successful strategies.

Third, Vanderschraaf and Skyrms (2003) provided a useful theoretical analysis of turn-taking in games. They pointed out that orthodox adaptive learning based on *fictitious play* (Brown, 1951) can never result in turn-taking. In fictitious play, a player chooses payoff-maximizing best replies to whatever the co-player's mixed strategy is believed to be, and the belief, based on empirical induction, is continually updated in the light of repeated rounds.

Vanderschraaf and Skyrms suggested modified learning dynamics that they called *Markov fictitious play* to explain adaptive learning of turn-taking, and they provided relevant data from computer simulations, but they did not specifically address the problem of the evolution of turn-taking through selection rather than learning.

Lastly, Helbing *et al.* (2005) studied adaptive learning in a specially devised traffic route-choice game in which it is evident that $2R < S + T$. After many rounds, pairs of players often established coherent oscillatory behaviour, and the authors explained this with a novel reinforcement learning model applicable to groups of arbitrary size, based partly on experience and partly trial-and-error behaviour. This provides further rigorous evidence of human turn-taking in an experimental game, complementing the work of Rapoport *et al.* (1976), reviewed above.

The relative neglect of cooperation through coordinated turn-taking can be attributed to an excessive focus by researchers on the Prisoner's Dilemma and related synchronization games. More attention needs to be devoted to alternation games and their role in the evolution of dynamic forms of cooperation, including turn-taking, about which much remains to be learned.

ACKNOWLEDGEMENTS

This research was supported by an Auber Bequest Award from the Royal Society of Edinburgh. The authors are grateful to Edmund Chattoe-Brown for his comments on an early version of the manuscript.

REFERENCES

Aumann, R. 1974. Subjectivity and correlation in randomized strategies. *J. Math. Econ.*, **1**: 67–96.

Aumann, R.J. 1987. Correlated equilibrium as an expression of Bayesian rationality. *Econometrica*, **55**: 1–18.

Axelrod, R. 1984. *The Evolution of Cooperation*. New York: Basic Books.

Axelrod, R. 1987. The evolution of strategies in the iterated Prisoner's Dilemma. In *Algorithms and Simulated Annealing* (L. Davis, ed.), pp. 32–41. Los Alamos, CA: Morgan Kaufman.

Axelrod, R. and Dion, D. 1988. The further evolution of cooperation. *Science*, **242**: 1385–1390.

Axelrod, R. and Hamilton, W.D. 1981. The evolution of cooperation. *Science*, **211**: 1390–1396.

Bendor, J., Kramer, R. and Stout, S. 1991. When in doubt: cooperation in a noisy Prisoner's Dilemma. *J. Conflict Res.*, **35**: 691–719.

Binmore, K. and Samuelson, L. 1992. Evolutionary stability in repeated games played by finite automata. *J. Econ. Theor.*, **57**: 278–305.

Blume, L.E. 2003. How noise matters. *Games Econ. Behav.*, **44**: 251–271.

Brown, G.W. 1951. Iterative solutions of games by fictitious play. In *Activity Analysis of Production and Allocation* (T.C. Koopmans, ed.), pp. 374–376. New York: Wiley.

Browning, L. and Colman, A.M. 2004. Evolution of coordinated alternating reciprocity in repeated dyadic games. *J. Theor. Biol.*, **229**: 549–557.

Croson, R.T.A. 1999. The disjunction effect and reason-based choice in games. *Organ. Behav. Hum. Decis. Proc.*, **80**: 118–133.

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.

Crown, C.L., Feldstein, S., Jasnow, M.D., Beebe, B. and Jaffe, J. 1993. The cross-modal coordination of interpersonal timing: six-week-olds infants' gaze with adults' vocal behavior. *J. Psycholinguistic Res.*, **31**: 1–23.

Duncan, S., Jr. 1972. Some signals and rules for taking speaking turns in conversations. *J. Pers. Soc. Psychol.*, **23**: 283–292.

Fogel, D.B. 1993. Evolving behaviors in the iterated Prisoner's Dilemma. *Evol. Comp.*, **1**: 77–97.

Frean, M.R. 1994. The Prisoner's Dilemma without synchrony. *Proc. R. Soc. Lond. B*, **257**: 75–79.

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.

Harsanyi, J.C. and Selten, R. 1988. *A General Theory of Equilibrium Selection in Games*. Cambridge, MA: MIT Press.

Helbing, D., Schönhof, M., Stark, H.-U. and Hoylst, J.A. 2005. How individuals learn to take turns: emergence of alternating cooperation in a congestion game and the Prisoner's Dilemma. *Adv. Complex Syst.*, **8**: 87–116.

Hofstadter, D.R. 1983. Computer tournaments of the Prisoner's Dilemma suggest how cooperation evolves. *Sci. Am.*, **248**(5): 14–20.

Holland, J.H. 1975. *Adaptation in Natural and Artificial Systems*, 2nd edn. Ann Arbor, MI: University of Michigan Press.

Kraines, D. and Kraines, V. 1989. Pavlov and the Prisoner's Dilemma. *Theory Decis.*, **26**: 47–79.

Kraines, D. and Kraines, V. 1995. Evolution of learning among Pavlov strategies in a competitive environment with noise. *J. Conflict Res.*, **39**: 439–466.

Lewis, D.K. 1969. *Convention: A Philosophical Study*. Cambridge, MA: Harvard University Press.

Lindgren, K. 1991. Evolutionary phenomena in simple dynamics. In *Artificial Life II: Studies in Complexity*, Vol. 10 (C.G. Langton, C. Taylor, J.D. Farmer and S. Rasmussen, eds.), 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. and Raiffa, H. 1957. *Games and Decisions: Introduction and Critical Survey*. New York: Wiley.

Manson, J.H., Navarrete, C.D., Silk, J.B. and Perry, S. 2004. Time-matched grooming in female primates? New analysis from two species. *Anim. Behav.*, **67**: 493–500.

May, R. 2006. Threats to tomorrow's world: Address of the President, Lord May of Oxford OM AC FRS, given at the anniversary meeting on 30 November 2005. *Notes Rec. R. Soc.*, **60**: 109–130.

Moore, E.F. 1957. Gedanken-experiments on sequential machines: automata studies. *Ann. Math. Stud.*, **34**: 129–153.

Mooring, M.S. and Hart, B.L. 1997. Reciprocal allogrooming in wild impala lambs. *Ethology*, **103**: 665–680.

Neill, D.B. 2003. Cooperation and coordination in the turn-taking dilemma. In *Theoretical Apsects* [sic] *of Rationality and Knowledge: Proceedings of the Ninth Conference* (M. Tennenholtz, ed.), pp. 231–244. New York: Association for Computing Machinery.

Nowak, M.A. and May, R.M. 1992. Evolutionary games and spatial chaos. *Nature*, **359**: 826–829.

Nowak, M.A. and Sigmund, K. 1992. Tit for tat in heterogeneous population. *Nature*, **355**: 250–253.

Nowak, M.A. and 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. and Sigmund, K. 1994. The alternating Prisoner's Dilemma. *J. Theor. Biol.*, **168**: 219–226.

Nowak, M.A. and Sigmund, K. 1998. Evolution of indirect reciprocity by image scoring. *Nature*, **393**: 573–577.

Nowak, M.A., May, R.M. and Sigmund, K. 1995. The arithmetics of mutual help. *Sci. Am.*, **272**(6): 50–55.

O'Connell, D., Kowal, S. and Kaltenbacher, E. 1990. Turn-taking: a critical analysis of the research tradition. *J. Psycholinguistic Res.*, **19**: 345–373.

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. and Budescu, D.V. 1997. Randomization in individual choice behavior. *Psychol. Rev.*, **104**: 603–617.

Rapoport, A. and Chammah, A.M. 1965. *Prisoner's Dilemma*. Ann Arbor, MI: University of Michigan Press.

Rapoport, A., Guyer, M.J. and Gordon, D.G. 1976. *The 2 × 2 Game*. Ann Arbor, MI: University of Michigan Press.

Rosenthal, R.W., Shachat, J. and Walker, M. 2003. Hide and seek in Arizona. *Int. J. Game Theory*, **32**: 273–293.

Roth, A. and Erev, I. 1995. Learning in extensive-form games: experimental data and simple dynamic models in the intermediate term. *Games Econ. Behav.*, **8**: 164–212.

Rousseau, J.-J. 1755. *Discours sur l'Origine et les Fondements de l'Inegalité Parmi les Hommes* [Discourse on the Origin and Foundations of Inequality Among Men]. In *J.-J. Rousseau, Oeuvres Complètes*, Vol. 3. Edition Pléiade, Paris.

Sheposh, J.P. and Gallo, P.S. 1973. Asymmetry of payoff structure and cooperative behavior in the Prisoner's Dilemma game. *J. Conflict Res.*, **17**: 321–333.

Trivelpiece, W.Z., Trivelpiece, S.G. and Volkman, N.J. 1987. Ecological segregation of Adelie, Gentoo, and Chinstrap Penguins at King George Island, Antarctica. *Ecology*, **68**: 351–361.

Trivers, R.L. 1971. The evolution of reciprocal altruism. *Q. Rev. Biol.*, **46**: 35–57.

Vanderschraaf, P. and Skyrms, B. 2003. Learning to take turns. *Erkenntnis*, **59**: 311–348.

Weimerskirch, H., Stahl, J.C. and Jouventin, P. 1992. The breeding biology and population dynamics of King Penguins *Aptenodytes patagonica* on the Crozet Islands. *Ibis*, **134**: 107–117.

Worden, L. and Levin, S.A. 2007. Evolutionary escape from the Prisoner's Dilemma. *J. Theor. Biol.*, **245**: 411–422.

Wu, J. and Axelrod, R. 1995. How to cope with noise in the iterated Prisoner's Dilemma. *J. Conflict Res.*, **39**: 183–189.

Young, H.P. 1998. *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*. Princeton, NJ: Princeton University Press.