J. Genet., Vol. 66, No. 1, April 1987, pp. 69–84. © Printed in India.

Evolution of cooperation by reciprocation within structured demes

N V JOSHI

Centre for Theoretical Studies, Indian Institute of Science, Bangalore 560 012, India

MS received 2 February 1987

Abstract. The iterative two-person Prisoners' Dilemma game has been generalised to the N-person case. The evolution of cooperation is explored by matching the Tit For Tat (TFT) strategy (Axelrod and Hamilton 1981) against the selfish strategy. Extension of TFT to N-person situations yields a graded set of strategies from the softest TFT, which continues cooperation even if only one of the opponents reciprocates it, to the hardest, which would do so only when all the remaining opponents cooperate.

The hardest TFT can go to fixation against the selfish strategy provided it crosses a threshold frequency p_c. All the other TFT are invadable by the selfish (D) or the pure defector strategy, while none can invade D. Yet, provided a threshold p_c is crossed, they can coexist stably with D. As N, the size of the group increases, the threshold p_c also increases, indicating that the evolution of cooperation is more difficult for larger groups. Under certain conditions, only the soft TFT can coexist stably against the selfish strategy D, while the harder ones cannot. An interesting possibility of a complete takeover of the selfish population by successive invasions by harder and harder TFT strategies is also presented.

Keywords. Evolution of cooperation; group selection; reciprocation; game theory ; Tit For Tat.

1. Introduction

Explaining the evolution of cooperation within the framework of natural selection has been one of the challenging problems of evolutionary theory. If two individuals cooperate, each of them ought to be better off than otherwise. However, if one of them 'cheats', i.e., manages to get the benefit from the other without cooperating in turn, he is likely to be much better off. The cheaters are thus expected to be at a selective advantage. Trivers (1971) pointed out that this phenomenon may be modelled by the two-person Prisoner's Dilemma game. Given a choice between cooperation and cheating, it is always better to cheat since the cheater does better than the cooperator regardless of what his opponent chooses. Both the players thus opt for cheating and are consequently worse off than they would have been had they cooperated. Trivers also showed that if the same individuals interact repeatedly (iterated version of the Prisoner's Dilemma game) and can base their choices on the 'experience' gained in the previous encounters, cheating may not be the best strategy. He went on to show that a large degree of asymmetry between the benefits and costs associated with the acts of cooperation could lead to the evolution of reciprocal altruism.

A formal model for the evolution of cooperation based on reciprocal interactions has been investigated by Axelrod and Hamilton (1981). They consider the iterated version of the Prisoner's Dilemma game, and adopt a probabilistic treatment for the frequency of repeated interactions between the same pair of individuals. Under this scenario a rich variety of complex strategies is possible. From an analysis of these strategies they have been able to identify a set of robust and evolutionarily stable cooperative strategies.

In nature, however, an individual often interacts simultaneously with several individuals. It is therefore of interest to model such situations as N-person versions of the iterated two-person game. We therefore explore in this paper the outcomes of competition between the selfish strategy and the Tit For Tat (TFT) strategy, proposed by A Rapoport, which was highly successful in the computer tournament organised by Axelrod (1984). An extension of TFT to N-person situations leads to a set of graded strategies, and we find that some of these can coexist in stable equilibria with the selfish strategy.

The basic model including the payoff matrices and the dynamics of genotype frequencies is described in §2. The results of competition between two interacting strategies are described in §3, while §4 deals with situations involving more than two strategies at a time.

2. The model

2.1 *Payoffs to the players*

The two actions open to any player in a given game are to cooperate (C) or defect (D). For a two-person game, the four possible combinations of the choices by the two players are CC, CD, DC and DD where the successive letters indicate the choice made by the first and the second player, respectively. The same notation can be used to denote the payoff obtained by the first player as a result of the choices. Thus, if the outcome is CD, the payoff to the first player is also denoted by CD, and the value of the payoff to the second player DC. We assume that the elements of the payoff matrix satisfy the inequality $DC > CC > DD > CD$ as in the Prisoner's Dilemma.

In an N-person game, each participant faces $N-1$ opponents. Each can choose either C or D. The payoff to any player then depends on his strategy, as well as the strategies chosen by his $N-1$ opponents. Let n of the opponents of a player choose C and let $N-1-n$ choose D. We take the payoff to a player choosing C or D as $f(C)$ or $f(D)$, where

$$f(C) = [n.CC+(N-1-n).CD]/(N-1),$$
$$f(D) = [n.DC+(N-1-n).DD]/(N-1). \qquad (1)$$

In other words, we assume that the payoff accruing to a player is the average payoff the player would have received in a series of two-person games played with each of the opponents separately, the choices of strategies by each player remaining unaltered.

2.2 *The iterated Prisoner's Dilemma game*

In the interated version, the same participants may interact more than once. Under these conditions, the strategy for each game (interaction) could be specified in terms of the choices made by each player and his opponents in their previous interaction. The situations when players interact M times is customarily referred to as a game consisting of M moves.

It is assumed that the game played by each group consists of at least one move. In one version, each game consists of exactly M moves. Alternatively, the number of moves in the game played by each group could be a random variable. Following Axelrod and Hamilton (1981), the probability of continuation of the game is taken to be constant, denoted by w. Then the probability that the game consists of exactly K moves is given by

$$P(K) = w^{K-1}(1-w), \tag{2}$$

and M, the mean number of moves in a game is $1/w$.

2.3 *Dynamics of genotype frequencies*

Maynard Smith (1974) initiated the game theoretic approach to modelling competition between organisms; others have combined this with population genetics by equating payoffs with genetic fitnesses (Gadgil *et al* 1980; Hines 1980; Maynard Smith 1982). A similar approach is used here.

Consider an infinite population of asexual organisms with nonoverlapping generations, and with frequency p and $1-p$ of genotypes A and B, respectively. At the beginning of each generation, the population is subdivided into groups, with N individuals per group. Assuming random association, the frequency of groups containing exactly n individuals of genotype A is given by

$$F(N, n) = \binom{N}{n} p^n (1-p)^{N-n}, \tag{3}$$

where

$$\binom{N}{n} = \frac{N!}{(N-n)!\,n!}$$

Each genotype is assumed to code for a specific strategy. The payoff of an individual of genotype A, when it is in a group containing n opponents of type A, and when K moves are made, is denoted by $Y_A(K, n)$.

The average payoff of A is then given by

$$f(A) = \sum_{n=0}^{N-1} F(N-1, n) \sum_{K=1}^{\infty} P(K) Y_A(K, n), \tag{4}$$

where $P(K)$ is the probability that the N-tuple plays a game consisting of exactly K moves.

We assume that the average payoff of A measures the fitness of A. The mean fitness for the population is then seen to be

$$f = pf(A) + (1-p)f(B), \tag{5}$$

and the change in the frequency of A from the mth to $(m+1)$st generation is given by

$$p(m+1) = [f(A)/f]\ p(m). \tag{6}$$

All the offspring of a given generation mix and resettle in random associations of N individuals each and then repeat the cycle to give rise to the next generation.

Given the various payoffs associated with the strategies of the competing genotypes, one can thus study the changes in frequency of any genotype. This procedure can be readily generalised to a case where more than two genotypes interact.

2.4 *Elements of the payoff matrix*

As explained earlier (§2.1), if an individual chooses the strategy C, his payoff for that game would be a linear combination of CC and CD (depending on the strategies of the rest of the competitors in the group), while a choice of D would make it a linear combination of DC and DD. Hence, the average payoff of a genotype would be a linear combination of DC, CC, DD and CD. It can be seen that if $f(A)$ is greater than $f(B)$ for any value of frequency p, it will remain so if a constant term is added to each of the elements of the payoff matrix and/or if each is multiplied by a positive constant. Therefore, without loss of generality, we can take $DC = 1$, $CC = c$, $DD = s$ and $CD = 0$. Hence, the effects of variation in the values of the elements of the payoff matrix can be explored using only two parameters, viz., c and s.

2.5 *The strategies*

We would like to explore the success of various cooperative strategies against the pure defector strategy. As mentioned earlier, in a two-person game, the Tit For Tat strategy of A Rapoport proved to be extremely successful in the 'tournaments' studied by Axelrod (1984). Hence, we consider here possible generalizations of the TFT strategy to $N-$person games. The TFT strategy uses information about the choices made by the opponent in the previous move (one-step memory). For mathematical convenience, we exclude strategies which can exploit other kinds of information (the number of moves remaining to be played, for example).

In an $N-$person game, each individual has $N-1$ opponents. We define a TFT strategy of the type n $(0 \leq n \leq N-1)$ as follows: cooperate on the first move. From the next move onwards, continue to cooperate if at least n of the opponents have cooperated; if not, defect in the next move. Thus, $n = 0$ corresponds to the pure cooperative strategy. In the two-person game, n can be either 0 or 1, with $n = 1$ corresponding to the TFT strategy of Rapoport. In general, for an $N-$person game, there would be $N-1$ TFT strategies. The strategy with $n = N-1$ is the hardest and $n = 1$ the softest TFT. For other variants of TFT strategies, see Taylor (1975).

3. Competition between two interacting strategies

We consider first the competition between pure D and pure C strategists followed by competition between pure D and various TFT strategies. It will be seen that whether one considers the total number of moves in a game as fixed (at M), or whether they follow the distribution described by (2) of §2.2, the outcome is unchanged for a two-strategy interaction. However, these two lead to different outcomes when more than two strategies interact. This result, as well as some implications of assuming a fixed number of moves, would be discussed in §4.

3.1 *Competition between pure D and pure C*

Let the frequency of pure C in the population be p. The probability of a pure C finding itself in a group containing exactly n, C strategists amongst its opponents is $F(N-1,n)$, according to (3). With such a composition, payoff to C

$$= [n.c+(N-1-n).0]/(N-1)$$
$$= [n.c/(N-1)].$$

If the games played by such an N-tuple consists of K moves, the payoff is

$$Y_C(K,\ n) = K.n.c/(N-1). \qquad (7)$$

Hence, the average payoff of C in the population is

$$f(C) = \sum_{n=0}^{N-1} F(N-1,\ n) \sum_{K=0}^{\infty} P(K).\ Y_C(K,\ n)$$
$$= \sum_{n=0}^{N-1} \binom{N-1}{n} p^n(1-p)^{N-1-n} \sum_{K=0}^{\infty} w^{K-1}.(1-w)\frac{K.n.c}{N-1}\ , \qquad (8)$$

from (4), (2), (3) and (7). Regrouping the terms

$$f(C) = \frac{c}{N-1}\sum_{n=0}^{N-1} n\binom{N-1}{n} p^n(1-p)^{N-1-n}\left[\sum_{K=1}^{\infty} K w^{K-1}(1-w)\right]$$
$$= \frac{c}{N-1}(N-1)p.M.$$
$$= cpM \qquad (9)$$

The term in the square bracket is the average number of moves in the games played by an N-tuple. The payoff is thus seen to depend only on the mean value and not on the distribution of the number of moves in the games played.

A very similar calculation shows that for D,

$$Y_D(K,\ n) = K[1.n+(N-1-n).s]/(N-1)$$
$$= Ks+K.(1-s)/(N-1),$$

and

$$f(D) = pM(1-s)+Ms. \qquad (10)$$

Hence

$$f(D)-f(C) = p.M.(1-c)+(1-p).M.s\ . \qquad (11)$$

This will be > 0 for all values of p, indicating that D would continue to increase in the population for any value of p. Hence pure D goes to fixation when in competition with pure C.

3.2 *Competition between pure D and hardest* TFT

The hardest TFT strategist requires all the opponents to cooperate for it to continue cooperation in the subsequent moves. Hence, in all the N-tuples where at least one D is present, the TFT switch over to defection after cooperating in the first move. Only in the N-tuple with all TFT, C is chosen in every move.

If the frequency of TFT is p, it is readily seen that

$$f(D) = \sum_{n=0}^{N-1} F(N-1, n)\{[(1-w).(n.1+(N-1-n).s]\}/(N-1)$$

$$+ \sum_{K=2}^{\infty} K.P(K).s$$

$$= p(1-s)+M.s \; . \tag{12}$$

Interestingly, this is independent of N. For the TFT,

$$f(\text{TFT}) = \sum_{n=0}^{N-2} F(N-1, n)\left[\frac{(1-w).n.c}{N-1} + \sum_{K=2}^{\infty} K.P(K).s\right]$$

$$+ F(N-1, N-1) \sum_{K=2}^{\infty} K.c.P(K)$$

$$= p^{N-1}(M-1)(c-s)+p.c.+(M-1).s \; . \tag{13}$$

Comparison of (12) and (13) shows that at $p=0$, $f(D)>f(\text{TFT})$, i.e., D is uninvadable by TFT. On the other hand, at $p = 1$,

$$f(D)-f(\text{TFT}) = 1+(M-1).s-M.c \; .$$

If M, s and *c* have values such that $f(D)<f(C)$ then TFT is also uninvadable by D. In fact, the line

$$s = (M.c-1)/(M-1), \tag{14}$$

separates the region in (c,s) space where TFT is uninvadable (below the line) from the one where it is invadable (figure 1). With increasing M, a larger region in the (c,d) space becomes favourable to TFT. As expected, the minimum value of M needed for pure TFT to be uninvadable increases with s and decreases with increasing c.

For the range of parameters where TFT is uninvadable, there is a critical value of p given by

$$p_c^{N-1}(M-1)(c-s)-p_c(1-c-s)-s = 0, \tag{15}$$

where the fitness of the competing strategies is equal. Since $f(D)$ increases linearly with p, $f(C)$ increases monotonically with p, $f(D)>f(C)$ at $p = 0$ and $f(D)<f(C)$ at $p = 1$, there is only one solution to this equation in the range $0 = 1$. Moreover, this equilibrium between the two strategies is unstable. In other words, once TFT crosses the threshold frequency p_c (by various mechanisms, such as invasion by clusters, discussed by Axelrod and Hamilton 1981), it will go to fixation.

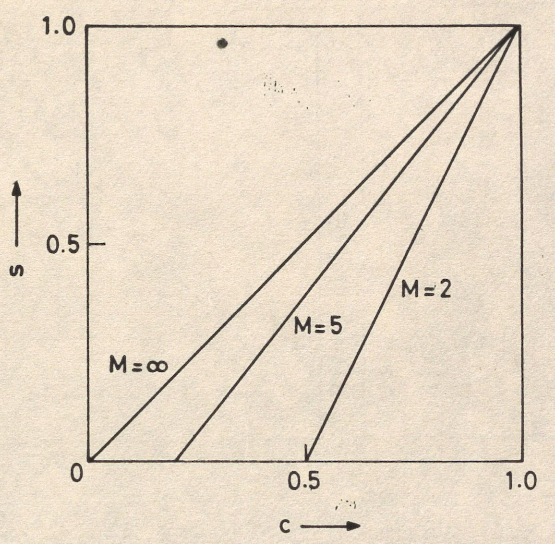

Figure 1. Regions in (c,s) space where the Hardest TFT is uninvadable by D. For a given value of M, the mean number of moves in the game, the region favourable to TFT lies below the corresponding line.

It is seen from (15) that p_c increases with s and decreases with increasing M and/or c as expected. Other parameters remaining constant, p_c increases with increasing N, suggesting that evolution of cooperation by this mechanism is more difficult in larger groups.

3.3 *Competition between pure D and soft* TFT

3.3a General formulation: Consider a soft TFT strategy S which continues to cooperate for the next move provided at least m of its opponents cooperate during the current move. The payoff of such a strategy in an N-tuple with n TFT strategists would be

$$\begin{aligned} Y_S(K, n) &= K.n.c/(N-1), \qquad n \geq m, \\ &= n.c/(N-1) + (K-1).s, \qquad n < m. \end{aligned} \tag{16}$$

The payoff of D in a similar situation is

$$\begin{aligned} Y_D(K, n) &= K.[n + (N-1-n).s]/(N-1), \qquad n \geq m, \\ &= [n + (N-1-n).s]/(N-1) + (K-1).s, \qquad n < m. \end{aligned} \tag{17}$$

Using these equations, the expressions for the average payoffs of D and S can be written down explicitly. Both $f(S)$ and $f(D)$ are polynomials in p (the frequency of S in the population), with positive coefficients. The curves $f(S)$ and $f(D)$ increase monotonically with p and are concave upwards (e.g. figure 2). When $p = 0$, both D and S find themselves with all the $N-1$ opponents as D. The average payoff of D in such a case is $M.s$, while that of S is $(M-1).s$ (since it cooperates in the first move); D is thus uninvadable by S. Near $p = 1$, each of them have $(N-1)$ opponents of the S type who cooperate for all the games. Hence,

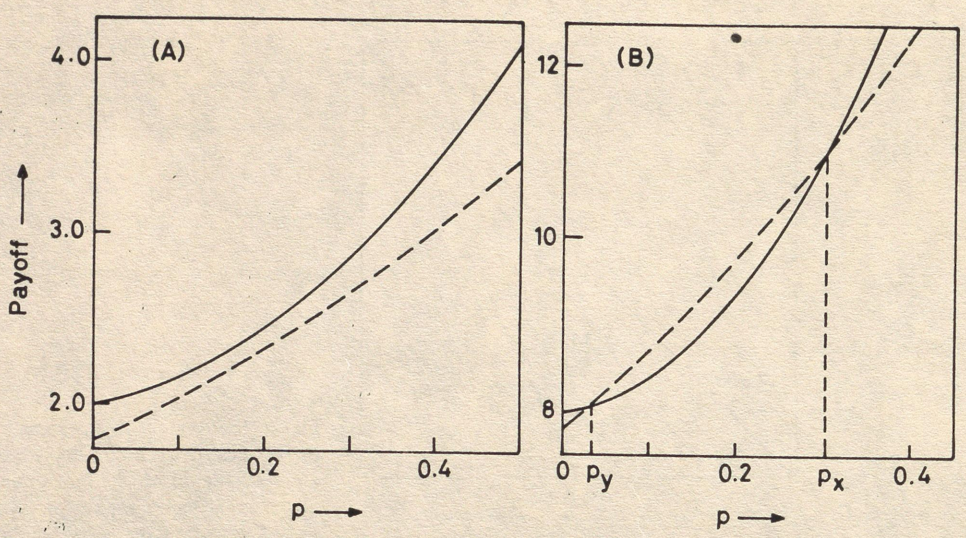

Figure 2. Payoffs to TFT (- - -) and D (——) as functions of p, the frequency of Soft TFT for $N = 3$, $c = 0{\cdot}6$ and $s = 0{\cdot}2$. A) For $M = 10$, D always goes to fixation. **B**) For $M = 40$, a stable coexistence between TFT and D is possible. The values of frequency p_y and p_x correspond to unstable and stable equilibrium, respectively.

$f(D) = M > f(S) = M.c$ and S is invadable by D (figure 2A). Under these circumstances, D is expected to go to fixation.

Interestingly, though $f(D) > f(S)$ at both $p = 0$ and $p = 1$, for some combination of the parameters M, s and c, it is possible for the fitness of S to be higher than that of D for some intermediate values of p (figure 2B). Under such circumstances, the curves $f(D)$ and $f(S)$ intersect at two points; at these values of p, the two strategies are in equilibrium. It is seen from figure 2B that the point corresponding to p_y is an unstable equilibrium, whereas the one corresponding to p_x is a stable one. In other words, once the strategy S is able to cross the threshold p_y, it would be able to coexist stably with D, at a frequency p_x.

3.3b *Soft* TFT *for* $N = 3$: For a three-person game, only one soft TFT is possible, viz., to continue cooperation even if only one of the other two opponents cooperates in the previous move. Using (16) and (17), and the procedure given in §3.3a, the average payoffs as a function of the frequency p of the TFT can be written as

$$f(D) = p^2(M-1)(1-s) + p(1-s) + Ms,$$

$$f(S) = p^2(M-1)s + p[Mc - 2(M-1)s] + (M-1)s. \tag{18}$$

The fitness of D is greater than that of TFT at both $p = 0$ and $p = 1$. For a given value of M, one can determine the combination of c and s such that the two curves $f(\text{TFT})$ and $f(D)$ touch at one point; the fitnesses of TFT and D are equal at that point whereas at any other point, $f(D)$ is greater than $f(\text{TFT})$. Any higher value of c or lower value of s then yields two values of p (p_y and p_x) where the fitnesses of the competing strategies become equal, and for $p_y < p < p_x$ we have $f(\text{TFT})$ greater than $f(D)$. Hence it is possible for the two strategies to coexist. Regions in (c, s) space

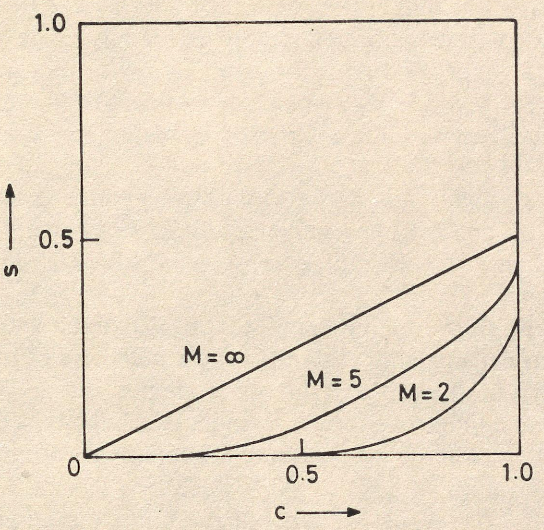

Figure 3. Regions in (c,s) space where a Soft TFT ($N = 3$) can coexist with D. For a given value of M, the region favourable to TFT lies below the corresponding curve.

where such a coexistence between S and D is possible are shown in figure 3 for a few values of M.

As expected, the region where coexistence is possible increases with increasing M. Low values of s and high values of c favour coexistence. A comparison of figures 1 and 3 is of interest. It is seen from figure 1 that as $M \rightarrow \infty$, in the entire region $c > s$, the hard TFT is uninvadable by D. On the other hand, as seen from figure 3, even as $M \rightarrow \infty$, the region favourable for a soft TFT is much smaller.

Is there, then, any advantage at all which a soft TFT enjoys over a hard TFT? Table 1 gives the values of threshold frequencies for various values of M for typical values of c and s. It is seen from the table that under some circumstances the threshold which a soft TFT needs to cross in order to establish itself in a stable coexistence with D is lower than the one needed for the hard TFT. Hence, the former is more likely to establish itself in a population though it would never go to fixation.

Table 1. Some values of s, c and M where a Soft TFT has a lower threshold frequency compared to the Hardest TFT.

N	c	s	M	Soft TFT	Threshold Soft TFT	Threshold Hardest TFT
3	0·85	0·25	5	$S(2,1)$	0·250	0·302
3	0·75	0·15	10	$S(2,1)$	0·040	0·176
4	0·95	0·15	5	$S(3,1)$	0·089	0·331
4	0·85	0·15	10	$S(3,1)$	0·048	0·287
4	0·85	0·25	5	$S(3,2)$	0·297	0·441
4	0·65	0·05	10	$S(3,2)$	0·092	0·294
5	0·65	0·05	5	$S(4,1)$	0·048	0·546
5	0·50	0·05	10	$S(4,1)$	0·032	0·513
5	0·95	0·15	5	$S(4,2)$	0·158	0·427
5	0·99	0·35	10	$S(4,2)$	0·272	0·433
5	0·75	0·25	5	$S(4,3)$	0·500	0·594
5	0·85	0·45	10	$S(4,3)$	0·472	0·532

3.3c *Soft* TFT *for* $N > 4$: Using the equations described in §3.3a. and the procedure outlined in §3.3b, one can obtain the regions of coexistence (allowed regions) in the (c, s) space for various soft TFT strategies for any value of N. Qualitatively, the results are as expected. As seen from figure 4, the allowed regions for the softest strategy for $N = 4$ is very small, for the less soft it is slightly larger, and so on. The allowed regions for all these strategies increase with increasing M.

An interesting result seen from figure 4 is that a softer strategy may be able to coexist with D, while the harder one may be unable to do so. For example, for $N = 4$, at $c = 0{\cdot}350$, $s = 0{\cdot}03$ and $M = 10$, $S(3,1)$ can coexist with D, while for $S(3,2)$ no coexistence is possible.

The advantage of the harder strategy lies in it being less prone to exploitation by D, while the disadvantage is in denying cooperation to some members of its own kind during the process; whereas for the softer of the two strategies, the situation is exactly the reverse. Over a limited range of parameter values, the balance seems to tilt in favour of the softer strategy. The regions where this happens are very small, and are characterized by low values of s.

It is possible to define a soft TFT not just in terms of the minimum number of cooperating opponents but in terms of the proportion of the cooperating opponents required by it for continuing cooperation. For example one can define a soft TFT which continues cooperation if at least half of its opponents cooperate, and defects otherwise. It is seen, however, that the regions for such a strategy in (c, s) space where it can coexist with D are not independent of N, but decrease with increasing N. Thus, the TFT strategy $S(10,5)$ which continued cooperation when at least 5 of its 10 opponents cooperated had a larger allowed region compared to, say, $S(20, 10)$.

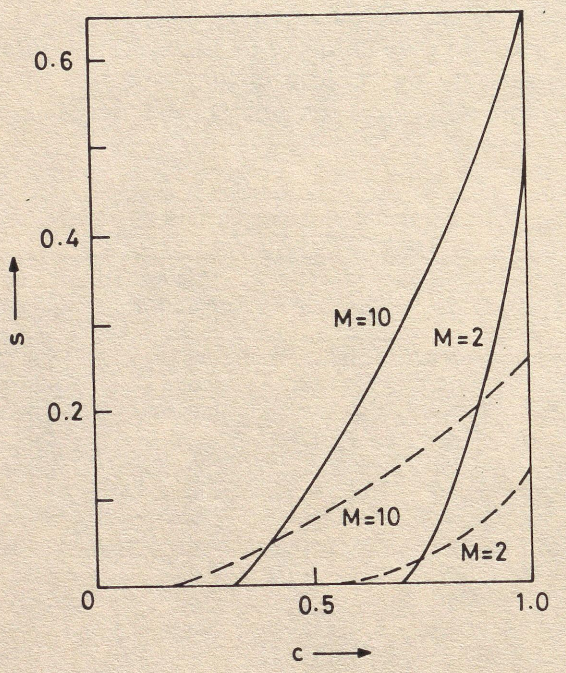

Figure 4. Regions in (c,s) space where Soft TFT can coexist with D, for $N = 4$, $M = 2$ and also $M = 10$. The region favourable to the Softest TFT lies below the dashed curves, while that for the harder strategy, lies below the solid curve.

To understand this result, we compare the payoff to a soft TFT $S(N, m)$ facing N opponents [n soft TFT and $(N-n)$ D's] with that to a D in a similar situation. If $n<m$, we get

$$Y_S(K,n) = n.c/N+(K-1)s < Y_D(K,n) = n.(1-s)/N+K.s,$$

and if $n>m$, still

$$Y_S(K,n) = K.n.c/N < Y_D(K,n) = [K.n.(1-s)/N]+K.s\ .$$

However, when $n = m$,

$$Y_S(K,m) = K.m.c/N,$$

while

$$Y_D(K,m) = [m(1-s)/N]+K.s.$$

If the values of K and c are high enough, and that of s low enough, it is possible for $Y_S(K,n)$ to be greater than $Y_D(K,n)$. Howwever, this is only a *necessary* condition. For $f(S)$ to be greater than $f(D)$, this advantage gained from the N-tuples with $n = m$ has to more than outweigh the disadvantage attained in the other N-tuples. For this to happen, p, the frequency of S, should be such that the frequency of N-tuples with $n = m$ is high, i.e., $F(N,m)$ should be maximized [(3)] with respect to p. It can be seen that the required value of p is equal to m/N.

However, the maximum value of $F(N,n)$ *decreases* with increasing N. For $F(N,N/2)$ for example, the maximum value of $F(6,3)$ is 0·3125, of $F(8,4)$ is 0·2734 and $F(10,5)$ is 0·2461. Hence, the allowed regions for soft TFT shrink rapidly as N increases, indicating once again that for large group sizes it is more difficult for cooperation to evolve.

4. Competition involving more than two strategies

The next logical step is to consider interactions where more than two strategies are involved. It is of interest to see whether the equilibrium between a soft TFT and D is invadable by a harder TFT. The motivation for such an analysis is provided from the following results seen for $N = 4$ for certain combinations of c, s and M: all the soft strategies can coexist with D. The softest of them, $S(3,1)$ has the lowest threshold. However, its equilibrium frequency is higher than the threshold of the next harder strategy, $S(3,2)$. Its equilibrium frequency, in turn, is higher than the threshold frequency of the hardest TFT strategy, which has the potential for taking over the population completely. It is tempting to explore whether a cascade process [$S(3,1)$ invading D, $S(3,2)$ invading $S(3,1)-D$ and $S(3,3)$ invading $S(3,2)-D$] can result in the complete elimination of D.

To start with, the simplest case ($N = 3$) is considered. The frequencies of the Soft TFT(S) and Hard TFT(H) are denoted by p_S and p_H respectively, and we investigate whether H can invade the S-D equilibrium. For it to be able to do so, its payoff at this equilibrium should be higher than that of S(or D). When invading an S-D equilibrium, H would face, in a triplet either SS, DD or SD as opponents. Now the payoff of H and S are identical against SS and DD opponents. However, payoff of H against SD is

$$= c/2 + P(2).(1+s)/2 + s. \sum_{K=3}^{\infty} P(K).(K-2)$$

$$= c/2 + [(1+s).(M-1)/2] + [(M-1)^2/M].s \ ,$$

while that of S against SD is simply $M.c/2$.

Thus, for H to successfully invade S-D equilibrium, c should be small and s should be large. These conditions, however, are exactly those which would render a stable equilibrium between D and S difficult.

The condition for equal fitness for H and S when H invades an S-D equilibrium is described by (using the above values of payoffs and simplifying)

$$s.(2M-1) - M.c + 1 = 0 \ . \qquad (19)$$

As seen in figure 5, this line cuts the $s = 0$ line at $1/M$, and $c = 1$ line at $(M-1)/(2M-1)$; and in the region above the line, invasion is possible. However, as seen from figure 3 [and as can be derived from (18)], these are exactly the points where the concave curve delineating the allowed region for S cuts the two lines $s = 0$ and $c = 1$; S can coexist with D *below* this curve. Therefore, H will never be able to invade the S-D equilibrium.

Consider now a situation where the total number of moves in the game is fixed at M. As Axelrod and Hamilton (1981) have shown, pure D is an evolutionarily stable strategy (ESS) under such a situation. To arrive at this result they have invoked

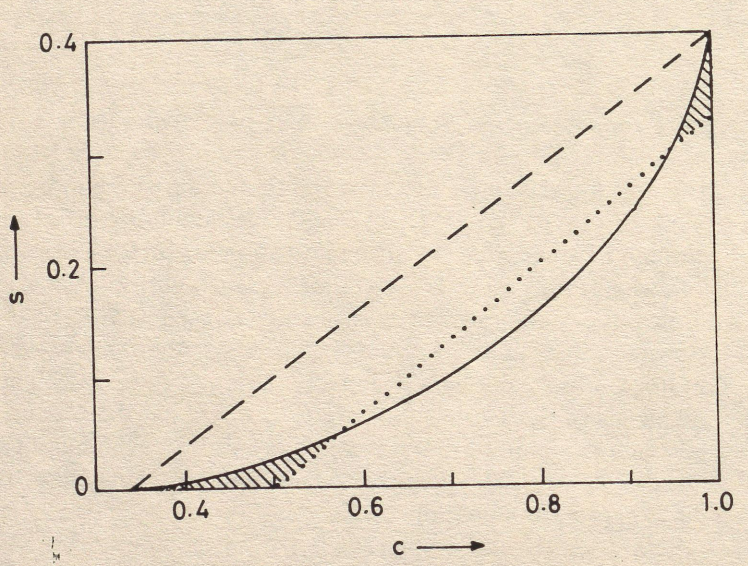

Figure 5. Invasion of the Soft TFT-D equilibrium by the Hardest TFT for N = 3, M = 3. A stable equilibrium between Soft TFT-D is possible only in the region below the curve. Invasion by the hardest TFT is possible only in the region above the dashed line if the number of moves in the game is geometrically distributed with mean M, and above the dotted line if the number of moves in the game is fixed at M. Accordingly, such an invasion is impossible for the former, while it is possible in the hatched regions for the latter.

strategies which involve a 'knowledge' of M, e.g., be a TFT upto $M-1$ moves and D at the last move, TFT upto $M-2$ moves and D for the last two moves etc. We, however, assume that the players have no 'knowledge' of M, and a strategy is defined only in terms of the choices made by the player and his opponents upto that move.

Thus when the games played by each triplet consists of exactly M moves, the payoff of H against SD becomes $c/2+(1+s)/2+(M-2).s$, while that of S against SD remains $Mc/2$. Hence, the condition for H to be able to invade S-D becomes

$$s.(2M-3)-(M-1).c+1>0. \tag{20}$$

As seen from figure 5, there are two small regions, one near $s = 0$ and the other near $c = 1$, where a coexistence between S and D is possible, and this equilibrium is invadable by H.

When such a situation occurs, under certain values of c, s and M, one can observe the following effect. The threshold needed for H to be able to invade D is higher than that needed by S. Therefore, S has a better chance of crossing the threshold, and reaching a stable equilibrium with D. Now, H can invade this combination and eventually eliminate D. The final composition of the population would be a mixture of S and H, depending on the size of the initial perturbation by H. In such a population, C, S and H are indistinguishable. One has thus seen evolution of cooperation taking place successfully against a selfish (D) strategy.

Analysis of an $N = 4$ case also reveals similar behaviour. When the number of games is distributed according to (2), no harder strategy can invade an equilibrium between D and the softer strategy. However, when the number of games is fixed, for some range of parameter values (a typical example being $M = 2$, $c = 0{\cdot}85$ and $s = 0{\cdot}05$), one can see the equilibrium between D and the Softest [$S(3,1)$] being invadable by the next harder strategy $S(3,2)$, and in turn, the $S(3,2)$-D equilibrium being invadable by the Hardest strategy, resulting in the complete elimination of D.

Such a takeover of a selfish population by successive invasions by gradually hardening TFT strategies may be of considerable interest. It must be emphasized, however, that the region in the parameter space where such behaviour is observed is extremely small.

5. Discussion and conclusions

Evolution of cooperation by reciprocation has been investigated in spatially structured populations. The choice of cooperation (or otherwise) by an individual is assumed to be influenced by the choices made by the other individuals in the group. Though 'strategy dependent selection' may be a better description for evolution occurring under this scenario, the framework of the N-person game theory is particularly suitable for analysing such situations (Charnov 1982; Maynard Smith 1982; Riechert and Hammerstein 1983). Gregariously living organisms exploiting a common resource–primate troops, breeding colonies of birds etc.–seem to be suitable systems for the applications of this model (Lombardo 1985).

We assume that the population is subdivided into groups containing N individuals each. The average fitness of an individual in groups with a higher proportion of TFT tends to be higher. Hence, following Wilson (1975, 1980) the evolution of cooperation in the present model can be described as an example of group selection. The differences in the average fitnesses of different groups are caused, however, by the reciprocal interactions between individual strategies.

Brown *et al* (1982) have presented a detailed analysis of the evolution of cooperation by reciprocation, using the TFT strategy. They have shown that as the ratio α/β increases (where α is the total number of interactions which an individual experiences in a generation and β is the number of interactions which are with individuals perceived as strangers), the threshold required by the TFT strategy for invading D becomes smaller. They have very ingeniously obtained values of α and β under a variety of conditions such as single partner/multiple partner models, finite/infinite memory, haploid/diploid organisms etc. and have also cast their model in a form such that their predictions can be compared to those made by Kin Selection theory. Another interesting and recent study of the evolution of cooperation (more specifically, helping behaviour) is by Peck and Feldman (1986). Using biologically plausible assumptions, they have shown that the threshold required by TFT for invading D can be made arbitrarily small. They have also investigated simultaneous competition involving C, D, and TFT, and have reported an interesting situation where invasion by D of a neutral equilibrium between C and TFT is initially successful, but leads eventually to fixation of TFT. Both these studies, however, have specifically considered situtions when population structure is absent. The present model, on the other hand, explores the evolution of cooperation in a structured population.

For two competing strategies A and B, if A is able to invade B, and is, in turn, uninvadable by B, it is generally taken to mean that A goes to fixation. We however show that where fitness has a nonlinear dependence on frequencies, it is still possible for the two strategies to coexist in stable equilibrium even if both the above conditions hold.

The present study has also shown that for a cooperative strategy to succeed against a selfish (defector) strategy, under certain conditions it is better to extend cooperation even though there is a risk of being exploited by a defector. On the other hand, under other conditions it is better to be more discriminating. We have also seen that a selfish population may be taken over by successive invasions by more and more discriminating strategies, though very discriminating strategy is unable to make any impact to start with against the selfish strategy.

It is of interest to examine the effects of relaxing some of the assumptions of the model presented in this paper. In diploid organisms with a one-locus-two-allele system, the two homozygotes TT and DD may be identified with TFT and D strategies. If T is recessive and if p is the frequency of T, then under Hardy-Wienberg equilibrium, the frequency of TFT strategists will be p^2. An examination of (15) indicates that the threshold frequency which a hard TFT has to cross in order to go to fixation is higher for diploids ($p_c^{1/2}$) compared to haploids (p_c) for a recessive gene. A similar result holds for a soft TFT [(18)]; however, the equilibrium frequency (coexistence with D) is also higher. If T is dominant, the thresholds are lower $[1-(1-p_c)^{1/2}]$, and the equilibrium frequency for a soft TFT is also lower. A further modification may be to consider the strategies as continuously varying traits coded for by a large number of genes. Aoki (1983, 1984) has analysed the evolution of TFT (and other) strategies in considerable detail, and has shown that biallelic models lead to predictions that are qualitatively and quantitatively different from those of polygenic models. A similar result is expected for the present model as well.

The assumption of random association in the formation of trait groups is also biologically unrealistic. Wilson (1980) has indicated that the variance between groups for most natural populations is more than that expected under a binomial distribution. Several mechanisms can lead to such an increase in the variance–assortative mating, kin recognition, high population viscosity etc. The net effect, however, is that the TFT are more likely to be associated with other TFT and the *D*'s with *D*'s. This would lower the threshold frequency for TFT's, indicating that the conditions for evolution of cooperation by reciprocation may be somewhat less stringent than predicted by the model. Finally, the assumption of infinite population size suppresses the effects of stochastic variations in the composition of groups. Studies with finite populations, which explore the effects of stochastic variations in group compositions and group sizes as well as in the elements of the payoff matrix are in progress.

The model considered here, albeit simple, has provided some insights into the different types of phenomena which may occur during evolution of cooperation. The Axelrod-Hamilton approach is being successfully used for analysing conflict situations in animals. Lombardo (1985) has described the mutual restraint in tree swallows as an example of the Tit For Tat strategy in an iterated Prisoner's Dilemma game. With more such studies forthcoming, it would be possible to verify the predictions of the present model.

Acknowledgements

It is a pleasure to thank Professors Madhav Gadgil and Sulochana Gadgil for introducing me to the subject, for many fruitful discussions, for encouragement and for numerous suggestions towards improvement of the manuscript. I also thank two anonymous reviewers for many helpful comments on an earlier version of the manuscript.

References

Aoki K 1983 A quantitative genetic model for reciprocal altruism: A condition for kin or group selection to prevail. *Proc. Natl. Acad. Sci. USA* 80: 4065–4068

Aoki K 1984 Quantitative genetic model of two policy games between relatives. *J. Theor. Biol.* 109: 111–126

Axelrod R 1984 *The evolution of cooperation* (New York: Basic Books)

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

Brown J L, Sanderson M J and Michod R E 1982 Evolution of social behaviour by reciprocation. *J. Theor. Biol.* 99: 319–339

Charnov E L 1982 *The theory of sex allocation* (New Jersey: Princeton Univ. Press)

Gadgil S, Nanjundiah V and Gadgil M 1980 On evolutionarily stable compositions of populations of interacting genotypes. *J. Theor. Biol.* 84: 737–760

Hines W G S 1980 An evolutionarily stable strategy model for randomly mating diploid populations. *J. Theor. Biol.* 87: 507–513

Lombardo M P 1985 Mutual restraint in Tree Swallows: A test of the TIT for TAT Model of Reciprocity. *Science* 227: 1363–1365

Maynard Smith J 1974 The theory of games and the evolution of animal conflicts. *J. Theor. Biol.* 47: 209–221

Maynard Smith J 1982 *Evolution and the theory of games* (New York: Cambridge Univ. Press)
Peck J R and Feldman M W 1986 The evolution of helping behaviour in large randomly mixed populations. *Am. Nat.* 127: 209–221
Riechert S E and Hammerstein P 1983 Game theory in the ecological context. *Annu. Rev. Ecol. Syst.* 14: 377–409
Taylor M 1975 *Anarchy and cooperation* (London and New York: Wiley)
Trivers R L 1971 The evolution of reciprocal altruism. *Q. Rev. Biol.* 46: 377–409
Wilson D S 1975 A theory of group selection. *Proc. Natl. Acad. Sci. USA* 72: 143–146
Wilson D S 1980 *The natural selection of populations and communities* (Menlo Park, CA: Benjamin-Cummings)

Reprinted from
The Journal of the Indian Institute of Science

J. Indian Inst. Sci., Nov.-Dec. 1987, **67**, pp. 521–533.

BOOK REVIEWS

Cellular automata machines edited by Tommaso Toffoli and Norman Margolus. The MIT Press, 55, Hayward Street, Cambridge, MA 02142, USA, 1987, pp. 259, $ 30. Indian orders to Affiliated East-West Press, Madras 600 010.

This book is presented in three parts, *viz*. Overview, Resources and Physical Modelling; each part consists of a number of rather short chapters. It gives an insight into cellular automata, where the primitive ingredients of the physical system that go into a model are reduced in to a single primitive ingredient, namely, the 'unit cell' governed by simple rules and coupled to identical cells by a uniform interconnection pattern.

In the first part, the authors have presented a historical review of Cellular Automata Machines (CAMs). The authors have also discussed CAM-6, the CAM developed at MIT, Massachusetts, which is the standard modelling environment used by the authors in developing physical models. Some rules pertaining to CAM-6 and a few trivial examples have also been presented in separate chapters. The four chapters in this part could have been merged into one single chapter without any discontinuity.

CAM rules are again discussed in the second part with some illustrative examples. The sources of information for computing the new state of the unit cell are discussed exhaustively in several chapters. Some of the methods for computing the new state and the related examples are explained in a rather circuitous manner. However, certain concepts like psuedo neighbours and Margolus neighbours have been presented lucidly. More significantly, the expressive power of these concepts as well as the occurrence of a random variable have been projected to indicate the direct mapping of cellular automata on to physical realizations and their suitability to a variety of modelling tasks.

The last part is devoted to express the role of cellular automata in physical modelling, which the authors claim to be complementary to that of differential equations. This claim is authenticated by some well illustrated models. To cite an example, the models on diffusion and equilibrium of gases and evolution of particle density in a gas discussed in the chapter on Diffusion and equilibrium. Similarly, the models on phenomena in hydrodynamics, Ising systems and optic waves underscore the power of cellular automata.

On the whole, the book is brought out in a systematic manner, the approach is good — the authors have started with the basic rules, modelling resources, then trivial modelling examples and have finally presented the nontrivial models. The book has a fair amount of programming examples in Forth, the language used for modelling physical phenomena in CAM-6. The CAM-6 software is supported on PC-DOS environment. The appendices in the book present a tutorial on Forth and the CAM architecture. The tutorial on Forth is just adequate to follow the examples given elsewhere in the book. Forth is a stack oriented language and the execution of individual processes is pipelined. The suitability of Forth for supporting parallel execution of processes in cellular automata is not very convincingly brought out by the authors.

The book can be recommended for researchers in computer architecture and physicists.

Department of Computer Science
and Automation
Indian Institute of Science
Bangalore 560 012.

L. M. Patnaik

Large scale scientific computing edited by P. Deuflhard and Engquist. Birkhauser Verlag, CH-4010, Basel, Switzerland, 1987, pp. 388, S. Fr. 58. Indian orders to Springer Book (India) Pvt Ltd., Panchasheel Park, New Delhi 110 017.

The book deals with computations involved for solving a variety of scientific and engineering problems. Mathematical models are used to describe the real life processes and numerical methods have been proposed to solve the associated large scale computations.

The book is a compilation of papers presented at a meeting on "Large scale scientific computing" held in July 1985. The book is organised into six parts with twenty-two chapters in all. The contents of the parts and chapters are briefly summarised below.

Part I covers initial value problems for ordinary differential equations (ODE) and initial boundary value problems (IBVP) for parabolic partial differential equations (PDE). Chapter I gives an introduction to numerical computations involved in semiconductor device modelling. The models used represent 3-D nonlinear PDEs which require enormous computations.

The second chapter is on the use of hierarchical bases in finite element computations, which are commonly used for 2-D elliptic PDEs. The new approach is promising for strongly coupled parabolic PDE systems. The next chapter reports about the extrapolation techniques developed for the numerical solution of quasilinear implicit ordinary differential equations arising in chemical reaction kinetics. The following chapter covers the combustion problem which requires the afore mentioned extrapolation techniques and multistep techniques. Physical considerations lead to the suggestion of a numerical scheme valid for large time scales, in the subsequent chapter. The last chapter in this part gives an algorithm for the numerical simulation of saturated–unsaturated flow through porous media.

Part II of the book covers boundary value problems for ODEs and elliptic PDEs. It includes five papers on the subject. The first one has an efficient path following technique applied to ODE-BVPs. Algorithmic details and numerical comparisons are the special features of this paper. The second paper is on the use of hierarchical basis techniques to reduce the computational effort of computing bifurcation diagrams for large non-linear parameter dependent systems. The next paper covers special ODE/PDE version of the pseudo-arc length continuation method. The procedures adopted to simplify the computations cannot adequately predict the detailed behaviour resulting from complex kinetics calculations. A subsequent paper by the authors promises to circumvent this drawback. The next paper has a hybrid numerical method for calculating complete theoretical seismograms. The last chapter in this part presents an efficient numerical algorithm for partial differential equations arising in 3-D geometries for viscous fluid flows.

Part III is on hyperbolic fluid dynamics. It has two articles. The first one is about a special class of shock capturing methods for the approximation of hyperbolic conservation laws. The method produces non-oscillatory solutions. Its special feature is that the computational stencil is always adapted to the solution. This results in a stability of the scheme. The second paper in this part addresses itself to a numerical method to solve compressible or incompressible Euler equation in 3-D using finite volume method with up to 6×10^5 grid cells.

Part IV shortly surveys some recent results for inverse problems in integral equations. The first two chapters deal with two alternative ways of treating the inverse radon transform problem arising in computer tomography. The third chapter presents numerical techniques for identification and optimal control in parabolic PDEs in a theoretical framework. The method, however, has to still address to some open questions.

Part V emphasises on large scale optimisation and optimal control problems. The first paper surveys the state-of-the-art solution techniques for solving large scale integer optimisation problems. MIMD architecture appears better suited problems than pipelined for vector machines for integer programming. Feedback control techniques for control and state constrained problems have been worked out in the setting of the multiple shooting method for ODE-BVPs in the next chapter. Two interesting applications have been presented. The feature meriting special mention is that the algorithm suggested is the only one capable of the treatment of control problems with control and state constraints. The last chapter of this part concerns the optimal control of a storage power plant.

The last part of the book has adaption of algorithms to supercomputers. This part has three chapters. Three main approaches used in the direct solution of sparse unsymmetric linear equations have been discussed in the first chapter. The next chapter presents a modification of adaptive gridding developed specially for supercomputers. Experience from the actual implementation of this concept is reported. The last chapter describes several approaches for adapting the existing numerical algorithms for use on supercomputers. A new iterative algorithm suiting vector computers has been developed.

Briefly, the book presents recent issues in large scale scientific computing covering initial and boundary value problems, inverse problems for integral equations and real life optimisation problems. The fields of application dealt with include semiconductor design, chemical combustion, seismology and fluid dynamics. The book is useful to scientists and engineers with interests in the computational aspects of ordinary and partial differential equations.

Department of Computer Science
and Automation
Indian Institute of Science
Bangalore 560 012.

L. M. Patnaik

AI in the 1980s and beyond–An MIT Survey edited by W. Eric. L. Grimson and Ramesh S. Patil. The MIT Press, 55, Hayward Street, Cambridge, Mass, 02142, USA, 1987, pp. 374, $ 24.95. Indian orders to Affiliated East-West Press, Madras 600 010.

This edited book brings together a collection of twelve papers by MIT staff members who participated in a conference in January 1986 entitled "Artificial intelligence: Current applications, trends and future opportunities". MIT has been a pioneering institution in AI since the birth of the field in the early 1960s. In the almost three decades of its existence AI has grown to be an evolutionary force in high technology industry and several laboratories in Japan, Europe and North America have joined the race. In India, 1987 has witnessed the establishment of a laboratory, Centre for Artificial Intelligence and Robotics (CAIR) in DRDO and the Department of Electronics has initiated the fifth generation computer project at five national institutions in India (TIFR, NCST, ISI, IISc and IITM). The MIT perspective is undoubtedly very important to the Indian R&D community in this frontier area of computer science.

The first paper by Winston is a brief introduction to the evolution of AI. He traces its career through its dawn in 1958, its early promises, the growth of LISP language, the emergence of time-sharing to meet the demands of AI researchers, its dark period in the late sixties and the early seventies and its renaissance in the 1980s with the commercial success of expert systems and the outgrowth of companies such as Symbolics Inc. and the LISP Machine Inc. out of MIT efforts. He discusses the current activities of the AI laboratory at MIT in the areas of robotics including perception, sensing, manipulation of objects and reasoning, language and learning and knowledge-based systems. Winston takes a working view regarding the primary goal of AI as making machines smarter and defines an intelligent robot as one which flexibly connects perception to action. He also raises several interesting questions for debate on expert systems, natural language interaction with machines and the role of robotics in increasing productivity in manufacturing. His main concern is with the commercial exploitation of AI technology.

In the second paper, Davis looks at knowledge-based systems at their stage of evolution in 1986. He notes that while rule-based and frame-based systems have been widely available, the problem of inexact reasoning has not been resolved satisfactorily. The central question is: "what makes a human expert an expert in his domain?" It is difficult to capture the human abilities of common sense reasoning, ability to reason both causally and qualitatively and the ability to learn from experience in the present generation of expert systems. Davis also notes the rapid commercialization of expert system technology and mentions systems such as PLANPOWER™ designed to help in individual financial planning. The present work at MIT is attempting to overcome the limitations of rule-based technology by exploring the theme of reasoning from first principles.

Szolovits explores the past, the present and the future of expert system tools and techniques. In view of the current revenues in the range of $ 100 m and market projections up to $ 800 m in 1990, an unlimited number of large system building tools and small PC-based tools are in the market. For the Indian readers, this article should provide lot of lessons about the over-blown commercialization and speculation in the field. The author's remarks about the underlying concepts in these together with the reading in between the lines should warn the Indian R&D laboratories not to be carried about by advertisements to buy these tools spending precious foreign exchange. (Indians have already spent enough by buying scores of books on expert systems which convey nothing but stereotypes of MYCIN). First generation expert systems can be designed and built in all good Indian laboratories as graduate or undergraduate projects!

The article by Patil focusses on AIM (Artificial Intelligence in Medicine). Since the phenomenal success of MYCIN at Stanford, medical diagnosis has always been treated as a prime application of expert systems. Patil describes the medical diagnosis problem vividly considering the issues of disease hierarchies, diagnostic reasoning with multiple disorders, etc. He describes the features of a program called ABEL whose knowledge base includes the shallow knowledge of associations between diseases and the deep pathophysiologic knowledge needed for accurate diagnosis. ABEL does more than conventional pattern classification by constructing a model that can explain the patient's illness. The paper ends with a speculation that expert systems might become common place in hospitals by the year 2000.

The paper of Rich and Waters on AI in software engineering is one pointing to a hope wherein real automatic programming may be feasible some day. Natural language understanding is one of the fond hopes of AI since the machine translation days. Berwick deals with intelligent natural language systems which might be able to reduce the imperfect present systems employing thousands of rules. The present effort at MIT aims at making these systems modular, more easily adaptable to future parallel processing architectures and more easily extendable to other dialects and languages. Zue looks at the related problem of automatic speech recognition and understanding. Over four decades of efforts have only produced isolated word recognition systems and recognition of continuous speech remains an elusive goal.

Brady is more optimistic about intelligent vision systems which have diverse applications. Vision poses a formidable challenge for AI presenting it a noisy and uncertain real world. The computational demands are enormous and it is even difficult to say what knowledge is needed for a robot to catch an object, a task which can easily be done by a little child. The entire SDI program depends on vision systems and the applications in defence are unlimited. Brady gives a very readable account of the diverse problems in vision and the trends in present research.

The last three articles are on robotics. The first of these deals with providing robots with sensing abilities and the second with robot programming. The article by Hollerbach deals with the comparatively mature field of robot hands. Designing multifingered hands places extreme demands on all aspects of robot technology. The final article by Brooks is on autonomous mobile robots which navigate around in environment.

In any book on topics at the forefront of contemporary AI research by a group of twelve authors, there are bound to be wide differences in scope and depth of treatment. This reviewer considers the chapters on medical diagnosis, natural language processing and vision to be particularly good. This book is recommended for all graduate students in AI for a quick and critical overview of the contemporary work in AI in a great institution. It reminds us, those connected with the *Journal of the Indian Institute of Science,* to use this as a medium for documenting the activities at IISc in particular, in a timely manner, to do at least a part of what the MIT Press is doing.

Department of Computer Science
and Automation
Indian Institute of Science
Bangalore 560 012.

V. V. S. Sarma

Systems that learn: An introduction to learning theory for cognitive and computer scientists by Daniel N. Osherson, Michael Stob and Scott Weinstein. The MIT Press, 55, Hayward Street, Cambridge, Mass. 02142, USA, 1986, pp. 205, $ 28.75. Indian orders to Affiliated East-West Press Pvt. Ltd., 6, Roselyn Gardens Apartments, 20/1A Barnaby Road, Madras 600 010.

There is a resurgence of interest in learning systems as there is a tremendous expansion in the field of artificial intelligence and as computers move towards the fifth generation. The book by Osherson *et al* is an attempt to build a learning theory from first principles.

Every theory is a product of its time. The present trend in cognitive science is to give a computer model for many phenomena. In keeping with this trend, the present book centres on the mathematical development of learning theory using fundamental concepts from computer science such as computable functions and recursion. One could contrast this approach with the classical work of Bush and Mosteller on stochastic models of learning which dealt with updating schemes for choice probabilities and was based on the theory of stochastic processes. The motivation for the present work comes from the acquisition of the first language by children and this theme is repeatedly used as an aid to intuition.

According to the paradigm of the book, learning typically involves 1. a learner, 2. a thing to be learned, 3. an environment in which the thing to be learned is exhibited to the learner, 4. the hypotheses that occur to the learner about the thing to be learned on the basis of the environment. Learning is said to be successful in a given environment if the learner's hypothesis about the thing to be learned becomes stable and accurate. On the basis of the above, learning theory is the study of systems that map evidence into hypotheses.

The book is divided into three parts. Part I advances a model of learning called identification suggested by Gold in 1967. Identification is intended as a model of language acquisition by children. However, the authors are conscious of several inadequacies of the model in explaining language acquisition. This is probably the reason why language acquisition is emphasized only as a motivating factor for the theory that is developed. It also makes some interesting speculations. One such speculation is that children may respond to a linguistic input not with one grammar but with a finite array of grammars each associated with some subjective probability.

Part II is devoted to a family of learning paradigms that results from modifying the definitions proper to identification. It examines various construals of "stability" and "accuracy" in the context of alternative criteria of successful learning.

Part III discusses efficient learning which has two demands. The learner must not examine too many inputs before settling for good on a correct hypothesis and second, the learner must not spend too long examining each input. This part also considers issues such as input required from the environment for learning and some aspects of probabilistic learning.

While the learning theory developed here is general, it is doubtful whether it can, in its present form,cover all aspects of learning. One weakness appears to be the nonutilization of the reaction of the environment for each hypothesis made by the learner. Such feedback exists in many situations in a probabilistic form, as for instance, in the two-armed bandit problem. Even in language acquisition, the child's parents often provide feedback in the form of corrections to the sentences formed by the child. Utilization of this type of feedback is extensively used by Bush and Mosteller and is further developed in models of learning automata. It appears that a much more powerful learning theory could be constructed by combining probabilistic ideas from learning automata and the computability theory of the present book. A glimpse of such a possibility is seen in the closing sections.

Department of Electrical Engineering
Indian Institute of Science
Bangalore 560 012.

M. A. L. THATHACHAR

Machine interpretation of line drawings by Kokichi Sugihara. The MIT Press, 55, Hayward Street, Cambridge, Massachusetts, 02142, USA, 1986, pp. 252, $ 30. Indian orders to Affiliated East-West Press Pvt. Ltd., 6, Roselyn Gardens Apartments, 20/1A Barnaby Road, Madras 600 010.

Understanding the shape of objects from two-dimensional line drawings is a problem that has received great attention. The importance of a good solution to this problem is apparent in the context of CAD/CAM applications. The 3-D shape data obtained by processing engineering drawings and hand-drawn line sketches of three-dimensional objects can then be further used in CAD/CAM systems towards meaningful applications. The book by Sugihara addresses itself exclusively to the issues in the specific problem of extracting 3-D shape of objects from a single 2–D line drawing and offers a reasonably exhaustive treatise on the topic put together under one cover for the first time. It consists of 11 chapters which are separately reviewed below.

The first introductory chapter starts on with an intuitive description of the problem of interpreting 3-D shape of polyhedral objects from 2-D line drawings. It then goes on to give a brief review of the literature on this and related topics. The aim of the book is stated quite unambiguously: "It should be said here that this book places emphasis on engineering rather than human science. Our aim is to construct a computational mechanism by which a computer can practically process line drawing data" (Chapter 1, page 5). In tune with the modular structure of the book, each subsequent chapter deals with one particular aspect of the whole problem. Chapter Two sets forth the assumptions which define the polyhedral and non-pathological nature of the scene whose 2-D line drawing can be meaningfully processed. Later, a set of simple procedural rules, for computing the set of all possible locally consistent edge labellings based on known types of physically possible trihedral vertices, is laid down. Finally, a pidgin Algol algorithm is presented which extracts, by constraint propagation for overall consistency, the small set of edge labelling schemes which are the final "Candidates for spatial interpretation" (Title of Chapter Two). Chapters Three and Four elaborate on discriminating between 'correct' line drawings which represent polyhedral scenes and 'incorect' ones which do not. Hidden line-eliminated drawings are considered in Chapter Three while Chapter Four deals with pictures in which hidden lines are also shown, as in engineering drawings. In both the cases, a necessary and sufficient condition for a correct interpretation of line drawings representing polyhedral scenes is shown to be equivalent to the existence of feasible solutions to a linear programming problem. Chapter Five consolidates the concept of algebraic structure of line drawings developed in the last two chapters and studies the nature and distribution of degrees of freedom in the choice of three-dimensional structure of the object represented in a line drawing. The superstrictness of the algebraic approach prevents it from being flexible such that it can tolerate slight vertex position errors in the line drawing due to digitization or hand-sketching approximation. The method of circumventing this superstrictness is studied in the following three chapters. Chapter Six describes detection of redundancy in the linear algebraic system of Chapters Three and Four. The method of deleting the redundant equations as well as automatically correcting the incorrectness in vertex positions is described in Chapter Seven. Chapter Eight presents an efficient algorithm to check whether an incidence structure (a set of ordered pairs of the form $<v_a,f_j>$ which represents the constraint that vertex v_a lies on face f_j) is generically reconstructible or not, in time proportional to the square of the number of incidence pairs. Chapters Nine and Ten describe how additional information, for example, specified lengths or angles (Chapter Nine), and surface texture and light intensity data (Chapter Ten) can be used to quantitatively fix the unique object shape most consistent with additional cues. This completes the methodological side of the line drawing interpretation problem. The last chapter (Chapter Eleven) is interesting but slightly out of context from the main theme.

It explores the correspondence between line drawings of polyhedral objects and rigidity of planar skeletal structures using a graph-theoretic approach.

There is hardly any noticeable typographical error in the book, but one finds a few minor syntactical errors, and on quite a few occasions, somewhat clumsy English sentence constructions are distracting. Figures are all quite clear, illustrative and unambiguous, except possibly Fig. 3.1 intended to show the orthographic projection of a polyhedral object. Any simple plan-elevation engineering drawing of a simple object would probably have been better illustrative. Mathematical treatments are terse, precise, and do not demand deeper background than basic linear algebra, elementary set theory and some knowledge of optimization techniques. A highly exhaustive set of 145 references is given.

Other than those involved in research in line drawing interpretation, readers interested in computer vision, graphics, artificial intelligence, robotics, and man-machine interaction in CAD systems involving 3-D objects will also find the book useful.

Department of Computer Science
and Automation
Indian Institute of Science
Bangalore 560 012.

L. M. PATNAIK

A vision of C and C: Computer and communications by Koji Kobayashi. The MIT Press, 55, Hayward Street, Cambridge, Mass. 02142, USA, 1986, pp. 190, $ 19.49. Indian orders to Affiliated East-West Press Pvt. Ltd., 6, Roselyn Gardens Apartments, 20/1A Barnaby Road, Madras 600 010.

This book is essentially a report on the various developments that have taken place in the fields of telecommunications, electronics and computers in Japan in the last about 60 years or so. The author of the book is an engineer turned executive, who started his career in 1929 and has grown with the technology. Hence one may expect the report to be authentic and accurate. The book also contains author's projection of how the field of computers and communications (C&C) is likely to evolve and affect the future of mankind.

The book begins with a summary of the pioneering events in Japan in the communications technology in the early years of development. The author springs a few surprises here claiming that many developments in this field took place in Japan much earlier than they did in the US, although most of the world is only aware of the US contribution. For example, the fundamental research results in switching theory published by C. E. Shannon in 1938 were apparently known in Japan by 1935 itself.

The author brings out that a single important factor that led to the tremendous growth in postwar Japan is the decontrol of the radio waves which were made available to the private sector in the early 1950s for radio and television broadcasting. As brought out by the author, it is interesting to learn that Japan had built its early computers using a device called parametron, which is a kind of resonance circuit, consisting of a small magnetic core, coils and capacitors, that works on the parametric excitation principle. Various developments and the process of coming together of the computers and communications have been well described by the author. He also presents well the social impact by expounding what he calls the concept of Man and C&C.

By the time one comes to the end of the book one feels a little tired of the claims made by the author. The book is so full of them that one wonders at the veracity of at least some of the statements. For example, the author claims to have envisioned the concept of integrating computers and communications, which he first presented in 1977. I quote here from a paper by R. M. Fano in 1972.

"The marriage of computers and communications has been celebrated and consummated. By now the honeymoon is over and the two partners are begining to face the realities of their interdependence".

If the reader is mentally prepared to put up with the many claims made, he would find in the book a very vivid account of the remarkable developments that have taken place in Japan in the field of C&C in the last about 60 years and that are likely to take place in the foreseeable future. The author's fond hope is that automatic real-time machine translation systems, that would enable people of different countries to converse freely with one another without having to use a common language, would become a reality before 2000 A.D. Let us too hope so.

Department of Electrical Communication Engineering and Computer Centre
Indian Institute of Science
Bangalore 560 012.

T. Viswanathan

Knowledge-based tutoring: The GUIDON program by William J. Clancey. The MIT Press, 55, Hayward Street, Cambridge, Massachusetts 02142, USA, 1987, pp. 377, $ 34.50. Indian orders to Affiliated East-West Press Pvt. Ltd., 6, Roselyn Gardens Apartment, 20/1A Barnaby Road, Madras 600 010.

This book describes a computer program called GUIDON that interacts with a student to teach him the knowledge needed for medical diagnosis problems. The specific field is that of diagnosing certain infectious diseases which is the field of expertise of MYCIN, the celebrated knowledge-based consultation program. The objective of GUIDON is to build an effective instructional tool utilising the knowledge base and problem-solving strategies of MYCIN.

There are many teaching programs developed which utilise an expert problem solver in the domains to teach the techniques to a student. But almost all such efforts are restricted to fields like geometry and algebra which have a precise formulation and well established norms as to what is the correct approach to a problem. As opposed to these, fields like medical diagnosis which are characterised by uncertain, heuristic knowledge and where there are no universal guidelines for evaluating problem-solving approaches, are much more difficult to handle. The GUIDON program described in this book is one of the first attempts at utilising a complex expert system at the heart of a teaching program. Since it is the first attempt, the results are not very exciting. But the book should still be of interest because it does provide some insights into this difficult teaching problem.

Briefly, the method followed by GUIDON is this: It maintains teaching knowledge distinctly from the domain knowledge MYCIN uses for actual problem solving. The program begins by getting MYCIN's solution on an example problem. This would have been obtained through MYCIN's backward chaining of rules. For GUIDON, MYCIN is modified to leave a detailed record of this problem-solving process. Then GUIDON constructs an AND/OR tree representation of MYCIN's solution which details the various rules tried, the failed subgoals, the data needed for determining the validity of different subgoals, etc. With this structure, GUIDON can rerun MYCIN's solution, but now in the forward direction, to determine which data is needed when, what inferences can be made with currently available data, etc. This AND/OR tree is what GUIDON uses to guide the student through the solution process. It gives the student the initial case data and then discusses with him various subgoals, waiting for him to ask for more data or seek help in solving some subgoals. The tutorial session is in mixed initiative mode, that is, either the teacher or the student can guide the conversation at any time. During the session, GUIDON has to develop and update a model of student knowledge to be able to interact with him intelligently. This is obtained by comparing the kind of data he is requesting and the type of inferences he is making against the same phase in MYCIN's solution. This process is fairly complex because the student need not know MYCIN rules in an identical fashion or even if he does, he need not apply them strictly in the same order as MYCIN does. All that can be expected is that given the same data, he should be able to make the same partial inferences as MYCIN does. For maintaining the student model, GUIDON uses its teaching knowledge.

The book systematically discusses the various problems involved in learning using the knowledge base of a typical expert system. The treatment is fairly general so that it is relevant not just for learning from MYCIN but to many other fields with the same general characteristics. Clancey begins by considering the issues involved in having the subject knowledge as a set of production rules. Chapters 2 and 3 describe the features of MYCIN rule set and what kind of information a teaching program should have about the structure of knowledge base for it to effectively compare student's solution with the expert's solution. The next two chapters address the problem of dialogue management especially in the case where the student maintains the initiative. Chapter 6 describes the teaching knowledge needed to construct and update the student model. The next two chapters contain the author's experiments with program and the extensions to the basic method. Chapter 9 deals with the conclusions of this effort. It contains a frank appraisal of the insights gained and the limitations of this program. A set of appendices (five in number) give details of the program and a full length example session.

GUIDON is an experimental system—a first attempt at trying to automate the teaching of a complex field. Since the program is no longer available, the interest in it is mainly of a theoretical nature. As the author himself points out, GUIDON did not answer any pedagogical questions about methods of teaching medicine. But it did demonstrate the usefulness of a specific representation technique to store the expert's solution in a complex domain. Also, it brought into focus limitations of MYCIN's representation of medical knowledge. Thus the book is of interest to anyone who is interested in computer-aided instruction and wants to know the problems involved when one steps out of simple well-structured domains and tries to teach a fairly complex subject.

Department of Electrical Engineering P. S. SASTRY
Indian Institute of Science
Bangalore 560 012.

Programming in FORTRAN by P. V. S. Rao. Tata McGraw-Hill Publishing Company Limited, 4/12, Asaf Ali Road, New Delhi 110 002, 1987, pp. 377, Rs. 45.

Considering the number of books already available on FORTRAN, it is not easy to justify writing yet another. *Programming in FORTRAN* by P. V. S. Rao, however, turns out to be a truly valuable addition to the textbooks on this subject.

The first two chapters lucidly explain the basic features of computers and, more importantly, the concept of algorithms. The detailed step-by-step illustrations should make it extremely easy for a student to understand the process of preparing flow charts.

The next six chapters describe the details of FORTRAN IV. It is the presentation in this section that makes the book not merely a good book on FORTRAN, but an excellent textbook *per se*. The description of each feature is very clear. Important points are repeated in separate paragraphs (liberally sprinkled with underlines) making them stand out from the main body of the text. Well annotated charts, which use a large sized typeface for FORTRAN statements, constitute yet another noteworthy feature, adding to the clarity as well as ease of learning.

Despite being an accomplished scientist and an expert in computer science, the author seems to have a keen insight into the kinds of difficulties and doubts the beginners are likely to be faced with. Thus, in the treatment of the ASSIGNED GOTO statement, the author points out how the reader may at first feel this to be an unnecessary feature, and follows up with an example where the use of an ASSIGNED GOTO statement is seen to improve the efficiency of the program.

The next chapter on 'Errors, efficiency and style' supplies many valuable hints and guidelines, which should go a long way towards making the student a competent programmer. The next chapter on features of FORTRAN-77 is also a very worthwhile addition.

Some of the detailed appendices make the book stand out from others on the subject. The descriptions of assembly language, process of compilation, etc., not normally found in elementary textbooks are concise, yet clear. The appendix on Numerical analysis, especially with its many solved examples, would undoubtedly make the student familiar with many techniques crucial for writing efficient and reliable programs for scientific and technical computations.

The numerous exercises at the end of each chapter are well designed to judge how well the student has understood the various aspects of every FORTRAN feature. Some of them are interesting enough to make writing programs for solving them a lot of fun. Hints and answers supplied at the end (though somewhat laconic) should make the book ideal for self-study as well.

Negative features, though not entirely absent, are few and far between. The book certainly deserved a better quality paper; the paper used is a little too thin for comfortable reading. The year of publication of the first edition is not indicated, and it is a little odd for a book published in 1987 not to cite any publication later than 1977. There are a few typographical errors and omissions. The style of writing is more pedantic than conversational. Neither of these, however, would seriously handicap the reader. A list of commonly available library functions would have been helpful.

In summary (to use one of the most popular current cliches used for describing software), this is an extremely 'User Friendly' book, and an ideal textbook on FORTRAN.

Centre for Ecological Sciences
Indian Institute of Science
Bangalore 560 012.

N. V. Joshi

Published by T.K.S. Iyengar, Executive Editor, Journal of the Indian Institute of Science, Bangalore 560 012.
Typeset and printed at Phoenix Printing Company Pvt Ltd. Attibele Indl Area, Bangalore 562 107.

Testing Sex Ratio Theory within Social Insects

Nirjanjan V. Joshi
Indian Institute of Science, Bangalore, India

In their pioneering paper of 1976 , Trivers and Hare applied sex ratio theory to test alternative hypotheses about evolution of sociality. Many factors governing optimal sex ratios in social insects have been investigated in detail (see Charnov , 1982 for a general review , Pamilo , 1982 for simulations of multilocus/multiallel control , Novacs , 1986 for a recent examination of empirical data). In the present work , effect of the simultaneous influence of local mate competition , finite brood size , male/female dimorphism and differential male/female dispersal on sex ratios is explored.

The Model

An infinite population with non-overlapping generations is assumed. On each of the single foundress nests , in addition to the worker brood , N reproductive offspring are produced , a fraction m being males (probability of an unfertilized egg being laid is m). Dimorphism is modelled by w such that one female is equivalent to w males. A fraction d_M of the males and d_F of females disperse (probability of dispersal is d_M and d_F respectively) to join the mating pool while sibmating takes place amongst those remaining at the nest. Singly inseminated females from the nests and the pool give rise to the next generation.

A single locus is assumed to control the sex ratio , the two alles acting additively. When a pure population of one allele is invaded by a small proportion of another allele coding for a different sex ratio , the dynamics of the allelic frequency change is expressed by a system of five linear difference equations. The elements of the corresponding matrix are functions of N , d_M , d_F , w and the sex ratios specified by the two alleles. If the dominant eigenvalue of the matrix is greater than 1 , invasion is possible. Optimal (evolutionarily stable) sex ratio is the one which cannot be invaded by any alternative sex ratio (for details , see Charnov 1982 , Joshi and Gadagkar 1985 , Joshi 1986).

Results

Under the joint influence of the two stochastic processes (dispersal and sex determination) , some of the females remain unmated on the nest and die without reproduction. The mean number of fertile females emerging

Eder/Rembold, *Chemistry and Biology of Social Insects*, Verlag J. Peperny, München 1987

per nest is

$$N (1 - m) [1 - (1 - d_F) (1 - m + m.d_M{}^{w})^{N-1}]$$

When the gyne controls the sex investment ratio , the optimal sex ratio $\hat{m}$ increases monotonically with d_F, reaching 1/2 at $d_F = 1$. The value of $\hat{m}$ decreases (female bias increases) with increasing N , increasing w and increasing d_M. For moderately high values of d_F and for $N > 50$, infinite brood approximation is quite satisfactory.

When the workers control the sex ratio , dependence of $\hat{m}$ on d_F is complicated. For high values of N and w and low values of d_M, $\hat{m}$ increses with d_F to 1/4 as expected. However , for a different set of values , it initially increases with d_F and then decreases. For the parameters (N=75, w=1.0 and d_M=0.90), $\hat{m}$ monotonically decreases with increasing d_F. This unexpected result (increased outbreeding leading to increased female bias) is traced to the complex interplay between genetic relatedness , stochasticity and local mate competition.

References

Charnov E.L., 1982. -- The theory of sex allocation. Princeton University press , Princeton , New Jersy.

Joshi N.V., Gadagkar R., 1985. -- Evolution of sex ratios in social hymenoptera : kin selection , local mate competition, polyandry and kin recognition. J.Genetics , 64, 41-58.

Joshi N.V., 1986. -- Evolution of sex ratios in social hymenoptera : consequences of finite brood size. J. Genetics (in press).

Nonacs P., 1986. -- Ant reproductive strategies and sex allocation theory. Q. Rev. Biol., 61, 1-21.

Pamilo P., 1982. -- Genetic evolution of sex ratios in eusocial hymenoptera : Allele frequency simulations. Am. Nat., 119, 638-656.

Trivers R., Hare H., 1976. -- Haplodiploidy and the evolution of social insects. Science, 191, 249-263