

The Failure of Success: Intrafamilial Exploitation in the Prisoner's Dilemma

Louis Marinoff

1. Introduction

A recent n -pair computer tournament for the repeated Prisoner's Dilemma (Marinoff 1992) amplifies and extends Axelrod's (1980a, 1980b) results, and demonstrates the relative robustness of co-operative maximization of expected utilities. For empirical and analytical purposes, the twenty competing strategies in the tournament are grouped into five "families," whose respective members share either common program structures, or similar conceptual functions. The five families are: the probabilistic family, the tit-for-tat family, the maximization family, the optimization family, and the hybrid family. Individual strategies or entire families can be selectively "bred" to exhibit (or to exclude) particular traits, or combinations of traits.

MAC, the most co-operatively-weighted member of the maximization family, is the most robust strategy in the tournament. *MAC* plays randomly during the first one hundred moves of each thousand-move game, with a co-operative weighting of $\frac{9}{10}$. It records the joint outcomes of the first hundred moves in an "event matrix," from which it computes its expected utility of co-operation (*EUC*), and of defection (*EUD*), from move 101 onward. The tournament payoff matrix, the maximizer's event matrix, and the expected utilities to which these matrices give rise are displayed in Figure 1.

MAC updates its event matrix after every move, and maximizes its expected utilities accordingly. That is to say, on its n^{th} move (where $100 < n \leq 1000$), *MAC* either co-operates or defects according to whichever of its expected utilities is the greater, based upon the previous $n-1$ moves.

MAC has three siblings in the tournament: *MAE*, *MEU*, and *MAD*. These siblings' program structures are identical to *MAC*'s, but their respective co-operative weightings for their hundred random moves are $\frac{5}{7}$, $\frac{1}{2}$, and $\frac{1}{10}$. A maximization family member's robustness in the tournament increases strictly with its initial co-operativeness. In the

		Column Player	
		<i>c</i>	<i>d</i>
Row Player	C	(3,3)	(0,5)
	D	(5,0)	(1,1)

C, c denote co-operation; *D, d* denote defection.

		Opponent	
		<i>c</i>	<i>d</i>
Maximizing Strategy	C	W	X
	D	Y	Z

$W = \text{frequency of } (C, c); X = \text{frequency of } (C, d);$
 $Y = \text{frequency of } (D, c); Z = \text{frequency of } (D, d);$
 expected utility of co-operation = $3W/(W + X)$
 expected utility of defection = $(5Y + Z)/(Y + Z)$

Figure 1: Prisoner's Dilemma payoff matrix and event matrix.

overall standings, out of twenty strategies, *MAC* placed first; *MAE*, third; *MEU*; eighth; *MAD*, tenth. Like its siblings, *MAC* is neither nice (where niceness means never defecting first) nor rude (where rudeness means always defecting first); rather, it is *nide* (where *nideness* means indeterminacy with respect to primacy of defection). And like its siblings, *MAC* is both provocable and exploitive. But unlike its siblings, *MAC* is initially co-operative enough to attain perpetual mutual co-operation with other provocable yet forgiving strategies, such as *TFT* (tit-for-tat). Thus *MAC* can become nice. These attributes, among others, account for *MAC*'s success in the environment under discussion.

The tournament reveals a number of interesting performance characteristics of the maximization strategies, but also exposes an ironic deficiency in their intrafamilial encounters: members of this family often fail to recognize one another, and their twins, in competition. While the random phase of its play permits a maximization strategy to "learn" about its opponent's responses by constructing an event matrix of joint outcomes, a pair of competing maximization strategies can "misconstrue" one another as random strategies, on the basis of their respective hundred-move event matrix profiles. Such cases of mistaken identity can result in perpetual mutual defection from move 101 onward. Pure defection is the optimal strategy against a random player, and as such is prescribed by the maximization calculus.

Figure 2 (reprinted from Marinoff 1992, p. 214) illustrates this general deficiency in intrafamilial maximization performance,¹ and also raises several perplexing questions. First, given that all members of this family experience sharp decreases between their average tournament scores and average intrafamilial scores, why is it that the magnitudes of the differences do not correspond to the order of increasing initial co-operativeness? From greatest to smallest, the strategic order of difference is: *MAC*, *MAD*, *MAE*, *MEU* (while the corresponding order of initial co-operativeness is *MAC*, *MAE*, *MEU*, *MAD*). Next, within the family itself, the order of success – *MAE*, *MEU*, *MAC*, *MAD* – is again altered with respect to initial co-operativeness. The second-most co-operative strategy finishes first within the family; the third-most co-operative strategy, second; the most co-operative strategy, third. Why does this unexpected order obtain? Only in competition against *MAD*, the least co-operative strategy, are the other strategies exploited in strict order of their increasing initial co-operativeness. Why? Finally, in competition against respective twins, the most successful pair is *MAE-MAE* (averaging 2594 points per game), followed by *MEU-MEU* (averaging 2384 points per game). But the *MAC-MAC* twins, which are weighted far more co-operatively than the others, average only 1807 points per game. Why does *MAC*'s overwhelming probability of co-operation during the first 100 moves ($\frac{9}{10}$, as opposed to $\frac{5}{7}$ for *MAE* and $\frac{1}{2}$ for *MEU*) result in a relatively poor performance between *MAC-MAC* twins? This is the most surprising and counter-intuitive result of the tournament.

2. Normally Distributed Scores

In order to appreciate what takes place when a maximization family member encounters a sibling, or a twin, one must recognize a strategic property peculiar to this family; namely, its members' sequential use of probabilistic, then deterministic algorithms. Thus, one observes two different phases in a maximization strategy's play: first, its construction of the initial event matrix for 100 moves; second, its calculation of expected utilities (and updating of the event matrix) for the subsequent 900 moves. But when maximization family members encounter

	<i>MAC</i>	<i>MAE</i>	<i>MEU</i>	<i>MAD</i>	Interfamilial Average	Tournament Average
<i>MAC</i>	1807	1849	1741	971	1592	2645
<i>MAE</i>	2123	2594	2356	987	2015	2503
<i>MEU</i>	1887	2396	2384	1003	1918	2362
<i>MAD</i>	1332	1266	1181	1029	1202	2086

Figure 2: The maximization family – interfamilial competition.

one another, their play takes on a reflected aspect, wherein certain symmetries, as well as asymmetries, become apparent.

One can identify five different kinds of algorithmic function, in the tournament environment: predetermined, purely probabilistic, purely deterministic, mixed probabilistic and deterministic, and sequential probabilistic and deterministic (see Marinoff 1992 for a detailed description of each strategic agent.) If two pre-determined and/or deterministic strategies are paired in a sequence of games, the scores of the given pair obviously do not vary from one game to another. If a probabilistic (or mixed probabilistic and deterministic) strategy is paired with any strategy *other than* a sequential strategy in a sequence of games, the scores of the given pair vary according to a normal distribution, in which the mean score approaches the most probable score as the number of games increases. When a maximization strategy meets a strategy that uses a mixed probabilistic and deterministic algorithm, the former's scores tend to be highly concentrated; the latter's normally distributed. The maximization family members' scores against one another, however, are neither concentrated nor distributed normally, with one noteworthy exception. In consequence, their average scores do not, as a rule, approach their most probable scores as the number of games increases.

Let the exception to the rule, which occurs in games involving *MAD*, be considered first. The extreme case of this exception obtains when *MAD* plays itself. Recall that during its first 100 moves, *MAD* co-operates randomly with probability $\frac{1}{10}$. Thus, the *a priori* probabilistic outcomes for a *MAD-MAD* pair are: $p(C,c) = \frac{1}{100}$; $p(C,d) = p(D,c) = \frac{9}{100}$; $p(D,d) = \frac{81}{100}$. So after 100 moves, the most probable event matrix contains entries $W = 1$; $X = Y = 9$; $Z = 81$, with associated expected utilities $EUC = 0.3$, $EUD = 1.4$, and the score tied at 129. The deterministic play that ensues from this matrix, from moves 101 to 1000, consists of 900 consecutive mutual defections. The game ends with the score tied at 1029. Since this score is a deterministic end-product of the most probable event matrix, it is the most probable score. Empirically, after five hundred games, *MAD's* average score was found to be 1029. The scores themselves appear to be distributed normally.

3. Non-Normally Distributed Scores

Next, consider an encounter between the *MEU-MEU* pair. Since *MEU* co-operates with probability $\frac{1}{2}$ during its first 100 moves, then the *MEU-MEU* pair has equiprobable outcomes during this phase: $p(C,c) = p(C,d) = p(D,c) = p(D,d) = \frac{1}{4}$. Thus, after 100 moves, the most probable event matrix has equal entries: $W = X = Y = Z = 25$, with associated

expected utilities $EUC = 1.499$, $EUD = 2.999$, and the score tied at 225. The deterministic phase of their encounter proceeds as follows. One-hundred-and-fifty consecutive mutual defections obtain between moves 101 to 250, with a concomitant steady decrease in the value of EUD .¹ By move 251, the value of EUD is driven below that of EUC , and 750 consecutive mutual co-operations ensue. After 1000 moves, the score is tied at 2625. Again, it is the most probable score.

Empirically, however, after 500 games of MEU versus MEU , the average score is found to be 2384. This is substantially less than the most probable value. The cause of the discrepancy is revealed in a histogram showing the distribution of scores for 500 games of MEU versus MEU . Figure 3 displays a non-normal distribution, with a minor prominence in the 1100–1200 point range, and a skewed distribution across the middle and upper ranges. The peak of the skewed distribution indeed coincides with the *a priori* most probable score, in the 2600–2700 point range. But the minor feature at the low end of the range, along with the overall skewness, diminishes the average score.

Next, consider an encounter between the MAE - MAE pair. Since MAE co-operates with probability $\frac{5}{7}$ during its first 100 moves, the MAE - MAE pairs' *a priori* probabilistic outcomes are: $p(C,c) = \frac{25}{49}$; $p(C,d) = p(D,c) = \frac{10}{49}$; $p(D,d) = \frac{4}{49}$. Thus, after 100 moves, the most

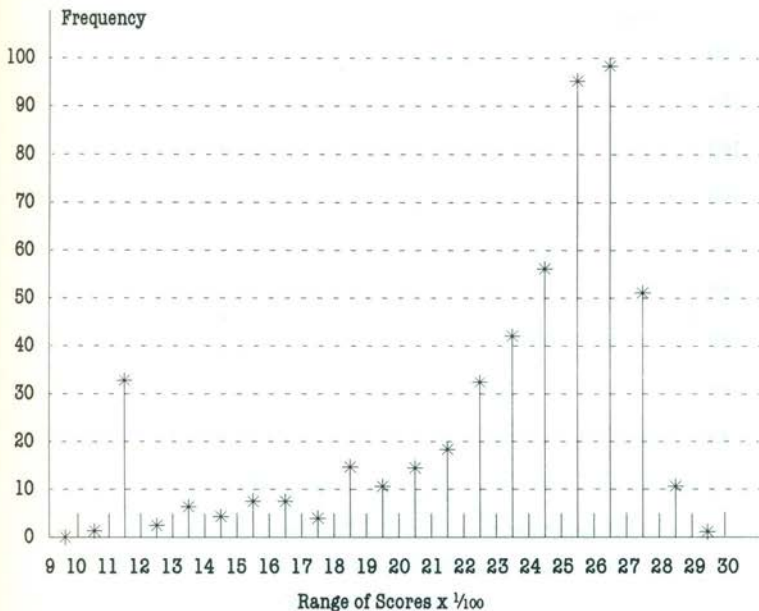


Figure 3: MEU versus MEU . Histogram of scores for 500 games. MEU 's average score: 2384.

probable event matrix has these entries: $W = 50$; $X = Y = 20$; $Z = 8$. The associated expected utilities are $EUC = 2.159$, $EUD = 3.879$, and the score is tied at 256. The deterministic phase of their encounter proceeds as follows. Forty-one mutual defections take place between moves 101 and 141, followed by 859 mutual co-operations. After 1000 moves, the score is tied at 2882 points. Again, this represents the most probable score.

Empirically, however, after 500 games of *MAE* versus *MAE*, the average score is found to be 2594 points. Again, a histogram (Figure 4) reveals the cause of the discrepancy between the most probable and the average scores. Figure 4 displays a non-normal distribution. While the most frequent scores by far occur in the 2800–2900 point range, which is the range of the most probable score, the skew of the distribution towards the lower ranges diminishes the average score by some 250 points. Other features of increasing prominence appear in the 1900–2000, 2300–2400, and 2600–2700 point ranges.

Finally, consider an encounter between the *MAC-MAC* pair. Since *MAC* co-operates with probability $\frac{9}{10}$ during its first 100 moves, the *MAC-MAC* pairs' *a priori* probabilistic outcomes are: $p(C,c) = \frac{81}{100}$; $p(C,d) = p(D,c) = \frac{9}{100}$; $p(D,d) = \frac{1}{100}$. Thus, after 100 moves, the most probable event matrix has these entries: $W = 81$; $X = Y = 9$; $Z = 1$. The

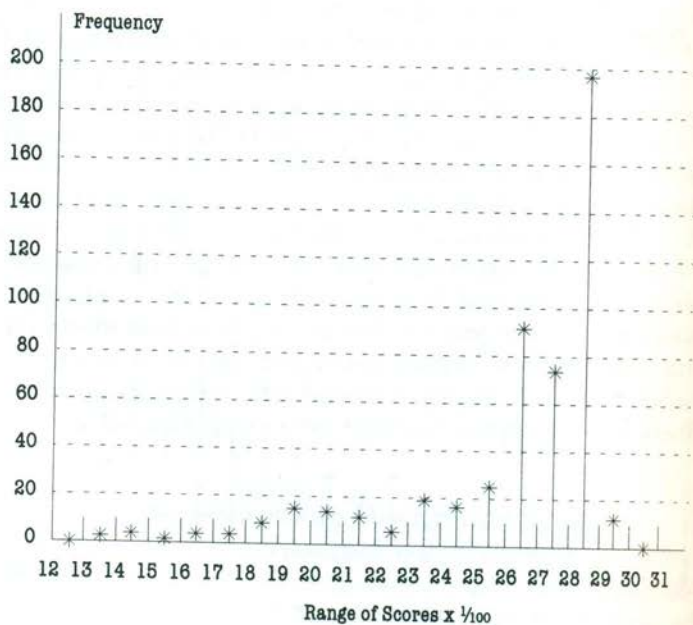


Figure 4: *MAE* versus *MAE*. Histogram of scores for 500 games. *MAE*'s average score: 2594.

associated expected utilities are $EUC = 2.699$, $EUD = 4.599$, and the score is tied at 288. In the deterministic phase, mutual co-operation commences on move 113, after only twelve consecutive mutual defections. The string of 888 prescribed mutual co-operations between moves 113 and 1000, in addition to the 81 probabilistic mutual co-operations during the first 100 moves, yields a total of 969 instances of mutual co-operation in a game of 1000 moves. The resultant score, which again represents the most probable score, is tied at 2965 points.

The competing *MAC-MAC* pair, however, realizes the largest empirical deviation in its family. After 500 games of *MAC* versus *MAC*, the average score is found to be 1807 points, a remarkable difference of 1158 points between the *a priori* most probable and *a posteriori* average scores. Again, a histogram (Figure 5) reveals the cause of this large discrepancy. Figure 5 shows a fragmented distribution of scores, with prominent features in the 1300–1400, 1600–1700, 2300–2400, and 2900–3000 point ranges. Empirically, the most probable score is 1300–1400 points. In addition, troughs appear between 1900–2200 and 2600–2700 points, from which ranges scores seem to be excluded. The histogram clearly illustrates how the average score for the *MAC-MAC* pair falls well below the most probable predicted score. But this illustration merely begs the question: Why does the distribution become so fragmented?

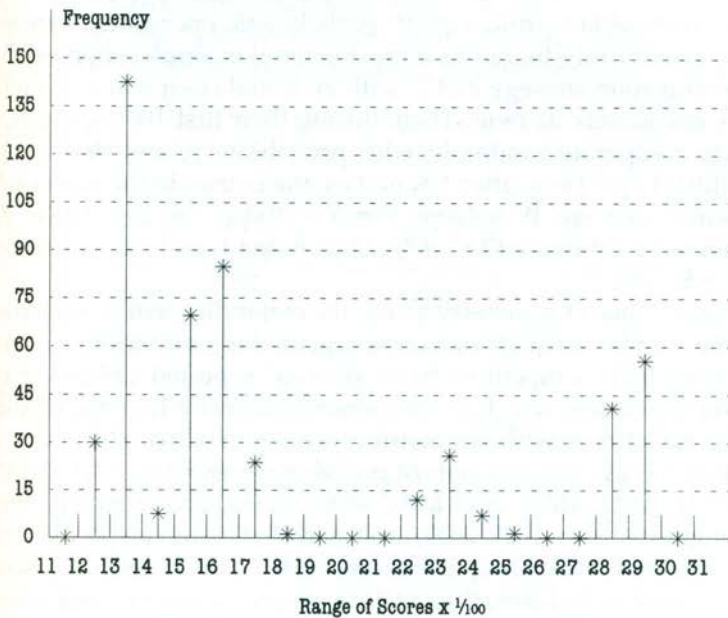


Figure 5: *MAC* versus *MAC*. Histogram of scores for 500 games. *MAC*'s average score: 1807.

Indeed, this is one of a number of questions raised by an examination of the distributions of scores between members of the maximization family. In the four cases considered, in increasing order of initial co-operative weighting, one finds: first, a concentration of scores at the low end of the scale; second, a skewed distribution with a minor prominence at the low end; third, a skewed distribution in the preliminary stages of fragmentation; and fourth, a fragmented distribution. One may wonder why these differences occur, given that each distribution represents a range of deterministic results stemming from a domain of probabilistic initial conditions. What causes such pronounced changes in the profiles of the distributions?

4. Analysis of Symmetric Event Matrices

Answers are found in an analysis of the event matrix itself. There are 176,851 different combinations of 100 trials of the four possible outcomes; in other words, for the first 100 moves in the iterated prisoner's dilemma, there are 176,851 possible event matrices. To facilitate analysis, one seeks to formulate a few general principles that extend to the many possible cases.

First, consider those matrices which are symmetric across their major diagonals; that is, event matrices in which the numbers of (C,d) and (D,c) outcomes are identical after 100 moves. As we have seen, such matrices obtain from *a priori* probabilistic encounters between maximization family twins. As a most general example, suppose that any maximization strategy *MAX*, with an initial co-operative weighting of p , encounters its twin. Then, during their first 100 moves, both strategies co-operate randomly with probability p , and defect with probability $(1-p)$. Thus, after 100 moves, the entries in the most probable event matrix are: $W = 100p^2$; $X = Y = 100p(1-p)$; $Z = 100(1-p)^2$. The expected utilities are $EUC = 3p$, $EUD = 4p+1$; and the score is tied at $100(1+3p-p^2)$.

The significance of symmetry across the major diagonal is as follows. When the number of (C,d) outcomes equals the number of (D,c) outcomes, then both competitors have identical expected utilities of co-operation and of defection. In consequence, from move 101 onward, their joint play is identical, with symmetric outcomes of either (D,d) or (C,c) .

In the *a priori* evaluations of most probable scores for the *MAD-MAD*, *MEU-MEU*, *MAE-MAE*, and *MAC-MAC* twins, one naturally finds increasing tied scores (1029, 2625, 2882, and 2965 points respectively) as the co-operative weighting increases. *MAD*'s most probable score against its twin is far lower than *MAD*'s siblings' most probable scores against their respective twins because, unlike *MAD*, the other siblings sooner or later attain mutual co-operation with their respective twins.

Empirically, it is found that the threshold co-operative weighting for the eventual attainment of mutual co-operation is $p = 37/100$ (in a game of 1000 moves with the payoffs of Figure 1). This is not a highly co-operative weighting; nevertheless, it does result in mutual co-operation from move 719 onward. The initial and final event matrices for this threshold weighting are displayed in Figure 6. Now, compare this result with that of a game in which the co-operative weighting of the competitors is $36/100$, or just below the threshold value, as displayed in Figure 7. While the initial conditions in Figures 6 and 7 scarcely differ, the final results admit of considerable difference.

Having established that the threshold weighting of $p = 37/100$ leads to the eventual attainment of mutual co-operation at move 719, one might next find the maximum rapidity with which such co-operation can be attained. The highest admissible value of p , to the nearest $1/100$, is $p = 99/100$. (If p equals unity, then EUD is undefined owing to division by zero.) After 100 moves, the event matrix for this maximum value of p contains entries $W = 98$; $X = Y = 1$; $Z = 0$. The expected utilities are $EUC = 2.97$, $EUD = 5$; the score is tied at 299. Two subsequent mutual defections, at moves 101 and 102, suffice to drive the value of EUD below that of EUC . Perpetual mutual co-operation ensues from move 103, with a resultant final score tied at 2995. (This is comparable to a final score between two nice strategies, which is tied at 3000.)

Evidently, for symmetric event matrices, the number of mutual defections required to bring on mutual co-operation can be represented

100 Moves:

		MAX	
		<i>c</i>	<i>d</i>
MAX	C	14	23
	D	23	40

$EUC = 1.14$; $EUD = 2.46$; score tied at 197

1000 Moves:

		MAX	
		<i>c</i>	<i>d</i>
MAX	C	296	23
	D	23	658

$EUC = 2.78$; $EUD = 1.14$; score tied at 1661

Figure 6: MAX versus MAX [$p(C) = p(c) = 37/100$], initial and final event matrices.

100 Moves:

		MAX	
		<i>c</i>	<i>d</i>
MAX	C	13	23
	D	23	41

$EUC = 1.08; EUD = 2.44; \text{score tied at } 195$

1000 Moves:

		MAX	
		<i>c</i>	<i>d</i>
MAX	C	13	23
	D	23	941

$EUC = 1.08; EUD = 2.09; \text{score tied at } 1095$

Figure 7: MAX versus MAX [$p(C) = p(c) = \frac{36}{100}$], initial and final event matrices.

as a decreasing exponential function of initial co-operative weighting. An exponential curve-fit yields the following equation:

$$n = f(p) = 7093e^{-7.164p}$$

where n is the number of mutual defections between move 101 and the onset of perpetual mutual co-operation and p ($\frac{37}{100} \leq p < 1$) is the co-operative weighting. The coefficient of determination for this exponential equation is 0.985.

Similarly, the final scores that result from these initial distributions can be fitted to a second exponential curve:

$$s = g(f(p)) = 3007e^{-.000958f(p)}$$

where s is the score after 1000 moves. The coefficient of determination for this expression is 0.9997.

Needless to say, while the numerical coefficients of both curves depend upon the particular payoff structure and the length of the game, the form of the curves is independent of these coefficients. In general then, both the play that ensues from event matrices exhibiting symmetry across their major diagonals, and the scores which result from this play, conform to simple mathematical expressions. This class of event matrix gives rise to regular and readily comprehensible outcomes.

5. Analysis of Asymmetric Event Matrices

That class of event matrices whose members do not exhibit symmetry across their main diagonals, is unfortunately (from the viewpoint of simplicity of analysis) the far larger of the two classes. The event matrices in this class give rise to the non-normal distributions displayed in Figures 3 through 5. It is possible (and desirable) to gain an understanding of why these distributions arise without having to analyze tens of thousands, nor even thousands of such matrices. Fortuitously, the process can be well represented by the tabling of results of a few dozen small probabilistic fluctuations about the most probable outcome, for each of the strategic pairs.

One first considers the case of *MEU* versus *MEU*, displayed in Figure 8. Recall the notation for entries in the generalized event matrix: W, X, Y , and Z are the respective numbers of (C,c) , (C,d) , (D,c) , and (D,d) outcomes. Columns labelled "Initial W,X,Y,Z " contain differing values of these variables after the first 100 moves, i.e., contain different probabilistically-generated event matrices. With each initial event matrix, described by a set $\{W,X,Y,Z\}$, is associated the move number on which perpetual mutual co-operation commences (column labelled "Perpetual (C,c) ") in the deterministic phase of the game (moves 101–1000) arising from that set. If no mutual co-operation occurs between moves 101–1000, the entry for that set reads "none." The column labelled "Final Score" associates the score (after 1000 moves) which results from the given initial set $\{W,X,Y,Z\}$.

The sets of $\{W,X,Y,Z\}$ values are arranged in blocks. Within each block, the values of W and Z are held constant, while the difference between X and Y increases. Each column of blocks holds the value of W constant, while the value of Z increases from block to block. Similarly, each row of blocks holds the value of Z constant, while the value of W increases from block to block. Thus, Figure 8 can be read both vertically and horizontally.

Reading down a column shows the effect of increasing initial difference in asymmetric outcomes (X,Y) within blocks, and of increasing initial mutual defections (Z) between blocks, upon the attainment of perpetual mutual co-operation and upon the final score. Reading across a row shows the effect of increasing the number of initial mutual co-operations (W) upon the attainment of perpetual mutual co-operation and upon the final score, with the number of initial mutual defections (Z) held constant and the variance in difference between asymmetric $(X \text{ minus } Y)$ outcomes held to unity.

Recall that for *MEU* versus *MEU* the most probable $\{W,X,Y,Z\}$ is $\{25,25,25,25\}$. In Figure 8, the sets of initial event matrices are representative

Initial W,X,Y,Z	Perpetual (C,c)	Final Score	Initial W,X,Y,Z	Perpetual (C,c)	Final Score	Initial W,X,Y,Z	Perpetual (C,c)	Final Score
20,33,32,15	none	1139-1139	25,30,30,15	move 386	2370-2370	30,28,27,15	move 262	2622-2622
20,34,31,15	none	1138-1143	25,31,29,15	move 423	2299-2299	30,29,26,15	move 281	2587-2587
20,35,30,15	none	1137-1147	25,32,28,15	move 464	2220-2220	30,30,25,15	move 301	2550-2550
20,36,29,15	none	1136-1151	25,33,27,15	move 510	2131-2131	30,31,24,15	move 323	2509-2509
20,37,28,15	none	1131-1156	25,34,26,15	move 562	2030-2030	30,32,23,15	move 347	2464-2464
20,30,30,20	move 651	1830-1830	25,28,27,20	move 324	2488-2488	30,25,25,20	move 213	2709-2709
20,31,29,20	move 755	1625-1625	25,29,26,20	move 354	2431-2431	30,26,24,20	move 229	2682-2682
20,32,28,20	move 885	1368-1368	25,30,25,20	move 386	2370-2370	30,27,23,20	move 245	2653-2653
20,33,27,20	none	1139-1139	25,31,24,20	move 423	2299-2299	30,28,22,20	move 262	2622-2622
20,34,26,20	none	1138-1143	25,32,23,20	move 464	2220-2220	30,29,21,20	move 281	2587-2587
20,28,27,25	move 497	2132-2132	25,25,25,25	move 251	2625-2625	30,23,22,25	move 185	2759-2759
20,29,26,25	move 567	1995-1995	25,26,24,25	move 273	2584-2584	30,24,21,25	move 199	2736-2736
20,30,25,25	move 651	1830-1830	25,27,23,25	move 298	2537-2537	30,25,20,25	move 214	2709-2709
20,31,24,25	move 755	1625-1625	25,28,22,25	move 324	2488-2488	30,26,19,25	move 229	2682-2682
20,32,23,25	move 885	1368-1368	25,29,21,25	move 354	2431-2431	30,27,18,25	move 245	2653-2653
20,25,25,30	move 346	2425-2425	25,23,22,30	move 213	2695-2695	30,20,20,30	move 151	2820-2820
20,26,24,30	move 389	2342-2342	25,24,21,30	move 231	2662-2662	30,21,19,30	move 162	2801-2801
20,27,23,30	move 439	2245-2245	25,25,20,30	move 251	2625-2625	30,22,18,30	move 168	2788-2788
20,28,22,30	move 497	2132-2132	25,26,19,30	move 273	2584-2584	30,23,17,30	move 186	2759-2759
20,29,21,30	move 567	1995-1995	25,27,18,30	move 298	2537-2537	30,24,16,30	move 199	2736-2736
20,23,22,35	move 277	2557-2557	25,20,20,35	move 166	2780-2780	30,18,17,35	move 127	2858-2863
20,24,21,35	move 309	2496-2496	25,21,19,35	move 181	2753-2753	30,19,16,35	move 127	2853-2868
20,25,20,35	move 346	2425-2425	25,22,18,35	move 196	2726-2726	30,20,15,35	move 126	2850-2875
20,26,19,35	move 389	2342-2342	25,23,17,35	move 213	2695-2695	30,21,14,35	move 126	2845-2880
20,27,18,35	move 439	2245-2245	25,24,16,35	move 231	2662-2662	30,22,13,35	move 125	2842-2887

Figure 8: MEU versus MEU, varying event matrices and scores.

of some probabilistic fluctuations in these values that would naturally occur in empirical trials. Three main tendencies, and one interesting exception to them, quickly become apparent.

First, within each block, the onset of perpetual mutual co-operation (when it occurs) is increasingly delayed by increases in the difference between X and Y . For a given number of mutual co-operations, a given number of mutual defections, and an initial unequal number of (C,d) and (D,c) outcomes, the *MEU-MEU* pair first proceeds to equalize the number of (C,d) and (D,c) outcomes. Once that happens, their expected utilities become equal, and the pair then defects until the value of *EUD* is driven below that of *EUC*. Perpetual mutual co-operation then ensues, and a tied final score results.³ The greater the initial difference between X and Y , the greater number of moves are required for their equalization, and the still greater number of moves must be made before mutual co-operation is attained. Thus, for a given W and Z , the smaller the initial difference between X and Y , the larger the final score.

Second, reading down the columns, one perceives that for a constant value of W , the onset of perpetual mutual co-operation is actually hastened as the initial number of mutual defections increases. Within certain probabilistic limits, which vary according to their initial weightings, the maximization strategies demonstrate the capacity of enlisting mutual defections in the service of perpetual mutual co-operation. While one wishes to refrain from lapsing into trite moralization, this counter-intuitive capacity suggests that, in certain instances, the game-theoretic end may justify the game-theoretic means.

Third, reading across the rows, one perceives that for a constant value of Z , the onset of perpetual mutual co-operation is hastened as the initial number of mutual co-operations increases. This tendency is not surprising, but reassuring in terms of the integrity of the maximization strategy.

In general, Figure 8 shows that perpetual mutual co-operation between *MEU-MEU* pairs, and thus their final scores, depend upon three factors. The scores tend to increase as W increases with Z fixed, as Z increases with W fixed, and as the difference between X and Y decreases with both W and Z fixed. One can amalgamate the first two tendencies, and observe that the final scores tend to increase as the sum of similar outcomes (W plus Z) increases; or, equivalently, as the sum of dissimilar outcomes (X plus Y) decreases. This observation, however, leads to the aforementioned exception.

The $\{30, X, Y, 35\}$ block boasts the largest W and Z values in Figure 8, yet the results that stem from this block are not altogether consistent with the tendencies so uniformly prevalent in the rest of the table. To begin with, the onset of perpetual mutual co-operation is hastened

(albeit only slightly) as the difference between X and Y increases, not decreases. And, as evidenced by the absence of tied final scores, the *MEU-MEU* pairs in this block attain perpetual mutual co-operation without having first equalized X and Y values, and without ever equalizing them. The scores themselves are the highest in the table, in keeping with this block's highest $W + Z$ sum. The significance of this unusual block will be brought to light in subsequent tables.

Meanwhile, Figure 8 does indeed account for the distribution of scores in Figure 3. One can observe the contributions towards skewness, with a majority of scores occurring in the 2400–2700 point range, and none exceeding 2900 points. Contributions to the minor prominence in the 1100–1200 point range occur when the sum of W plus Z falls below a certain threshold, making mutual co-operation unattainable within 1000 moves; or when the sum of W plus Z is theoretically sufficient for perpetual mutual co-operation, but the difference between X and Y is large enough to prevent its onset. These latter conditions prevail in the $\{20, X, Y, 15\}$ and $\{20, X, Y, 20\}$ blocks, respectively.

Next, a similar table is generated for *MAE* versus *MAE*. Recall that the most probable $\{W, X, Y, Z\}$ for the *MAE-MAE* pair is $\{52, 20, 20, 8\}$. Figure 9 displays corresponding fluctuations about these most probable values, and the results to which they give rise.

Reading down the first column of Figure 9, one observes that the two previous tendencies hold until the $\{40, X, Y, 14\}$ block; that is, the onset of perpetual mutual co-operation is hastened as the difference X minus Y decreases within blocks, and as the sum X plus Y decreases between blocks. The $\{40, 23, 23, 14\}$ matrix of the $\{40, X, Y, 14\}$ block also conforms to these tendencies. But the other matrices in that block yield results comparable to those of the $\{30, X, Y, 35\}$ block in Figure 8; that is, they give rise to perpetual mutual co-operation without first equalizing X and Y values, and the onset of mutual co-operation is hastened slightly as the difference X minus Y increases.

Reading down the second column, one observes that this departure from precedent tendency now becomes the norm itself. With the obvious exception of matrices in which X equals Y initially, the second column of blocks behaves as the last block in the first column. Note that, within each block except the first, the order of the onset of perpetual mutual co-operation is increasingly jumbled.

The most important overall effect of this departure, exemplified in the first three blocks of column two, is reflected in the final scores. Because the X and Y values are not equalized prior to perpetual mutual co-operation, the gap between the final scores increases as the initial difference between X and Y increases. Owing to the vicissitudes of chance during the first 100 moves, one member of the *MAE-MAE* pair

Initial W,X,Y,Z	Perpetual (C,c)	Final Score	Initial W,X,Y,Z	Perpetual (C,c)	Final Score	Initial W,X,Y,Z	Perpetual (C,c)	Final Score
40,29,29,2	move 227	2715-2715	50,24,24,2	move 169	2836-2836	60,19,19,2	move 140	2899-2899
40,30,28,2	move 239	2694-2694	50,25,23,2	move 203	2754-2809	60,20,18,2	move 257	2622-2622
40,31,27,2	move 252	2671-2671	50,26,22,2	move 206	2742-2812	60,21,17,2	move 623	2059-2059
40,32,26,2	move 265	2648-2648	50,27,21,2	move 209	2730-2815	60,22,16,2	move 668	1975-1975
40,33,25,2	move 280	2621-2621	50,28,20,2	move 211	2720-2820	60,23,15,2	move 717	1883-1883
40,28,27,5	move 216	2734-2734	50,23,22,5	move 158	2851-2856	60,18,17,5	move 623	2059-2059
40,29,26,5	move 227	2715-2715	50,24,21,5	move 209	2730-2815	60,19,16,5	move 668	1975-1975
40,30,25,5	move 239	2694-2694	50,25,20,5	move 211	2720-2820	60,20,15,5	move 717	1883-1883
40,31,24,5	move 252	2671-2671	50,26,19,5	move 210	2717-2827	60,21,14,5	move 263	2585-2850
40,32,23,5	move 265	2648-2648	50,27,18,5	move 212	2707-2832	60,22,13,5	move 859	1614-1614
40,26,26,8	move 195	2770-2770	50,21,21,8	move 148	2869-2869	60,16,16,8	move 124	2922-2922
40,27,25,8	move 205	2753-2753	50,22,20,8	move 211	2720-2820	60,17,15,8	move 717	1883-1883
40,28,24,8	move 216	2734-2734	50,23,19,8	move 210	2717-2827	60,18,14,8	move 263	2585-2850
40,29,23,8	move 227	2715-2715	50,24,18,8	move 212	2707-2832	60,19,13,8	move 859	1614-1614
40,30,25,8	move 239	2694-2694	50,25,17,8	move 214	2697-2837	60,20,12,8	move 265	2568-2868
40,25,24,11	move 180	2793-2798	50,20,19,11	move 210	2717-2827	60,15,14,11	move 263	2585-2850
40,26,23,11	move 195	2770-2770	50,21,18,11	move 212	2707-2832	60,16,13,11	move 859	1614-1614
40,27,22,11	move 205	2753-2753	50,22,17,11	move 214	2697-2837	60,17,12,11	move 265	2568-2868
40,28,21,11	move 216	2734-2734	50,23,16,11	move 443	2357-2357	60,18,11,11	move 268	2555-2875
40,29,20,11	move 227	2715-2715	50,24,15,11	move 213	2688-2853	60,19,10,11	none	1334-1359
40,23,23,14	move 140	2819-2819	50,18,18,14	move 129	2898-2898	60,13,13,14	move 110	2941-2941
40,24,22,14	move 166	2814-2824	50,19,17,14	move 214	2697-2837	60,14,12,14	move 265	2568-2868
40,25,21,14	move 166	2809-2929	50,20,16,14	move 443	2357-2357	60,15,11,14	move 268	2555-2875
40,26,20,14	move 165	2806-2836	50,21,15,14	move 213	2688-2853	60,16,10,14	none	1334-1359
40,27,19,14	move 165	2801-2841	50,22,14,14	move 523	2209-2209	60,17,9,14	none	1327-1372

Figure 9: MAE versus MAE, varying event matrices and scores.

finds that joint occurrences of its co-operation and its twin's defection outnumber joint occurrences of its defection and its twin's co-operation. In the $\{W,X,Y,Z\}$ region under consideration, this member's final score decreases, while its twin's increases, as the initial difference X minus Y becomes larger.

Then, suddenly, in the $\{50,X,Y,11\}$ block, a new phenomenon is manifest. Four of five sets in this block give rise to perpetual mutual co-operation between moves 210–214, with respective final scores within the 2688–2853 point range. But the $\{50,23,16,11\}$ matrix, which contains neither the largest nor the smallest (X,Y) difference in the block, gives rise to an unexpectedly large number of mutual defections, with the onset of perpetual mutual co-operation delayed until move 443. The resultant final score, tied at 2357 points, indicates that X and Y values are once again equalized during the game.

This phenomenon is increasingly more frequent, and more drastic, through the balance of column two, and throughout column three. For instance, consider what takes place in the $\{60,X,Y,8\}$ block. The first matrix, $\{60,16,16,8\}$, gives rise to early perpetual mutual co-operation, commencing on move 124, and the *MAE-MAE* twins attain a correspondingly high score, tied at 2922 points. But the second matrix, $\{60,17,15,8\}$, leads to comparative disaster: perpetual mutual co-operation does not commence until move 717, and the pair attains a correspondingly low final score, tied at 1883 points. Hence, a small increment in the difference between X and Y produces a momentous delay in the onset of perpetual mutual co-operation, with a correspondingly large decrement in the final scores.

The third matrix in the block, $\{60,18,14,8\}$, reverses the previous disaster. Perpetual mutual co-operation begins at move 263, which is now explicable in light of the initial (X,Y) difference. No equalization of (X,Y) values takes place, and the final scores are therefore fairly high but disparate, at 2585–2850 points. But the fourth matrix, $\{60,19,13,8\}$, leads to renewed disaster, with perpetual mutual co-operation commencing only on move 859, and a resultant low tied score of 1614 points.

The culmination of these alternating radical changes appears in the last two blocks of column three. The combination of a sufficiently large W plus Z sum and a sufficiently large X minus Y difference can result in perpetual mutual defection from move 101 to the end of the game. In such cases, the *MAE-MAE* pair attains scores of less than 1400 points.

Evidently, the event matrix becomes increasingly unstable as the sum of similar outcomes ($W + Z$) begins to exceed that of dissimilar outcomes ($X + Y$). The expected utilities associated with these outcomes begin to reverse their prescriptions with each increment of the (X,Y) difference, and the pendulum of joint outcomes swings steadily

away from perpetual mutual co-operation, and towards perpetual mutual defection, as W plus Z grows and X minus Y diminishes.

Figure 9 indeed accounts for the distribution of scores in Figure 3, albeit in an unexpected fashion. When random fluctuations about the most probable event matrix, $\{52,20,20,8\}$, are relatively small, the scores attained are fairly high. Larger fluctuations which *reduce* the sum $W + Z$ do not substantially reduce the final scores. But larger fluctuations which *increase* the sum $W + Z$ produce both the highest scores in the distribution (when X equals Y), as well as the lowest scores (when X minus Y is sufficiently large).

Next, a similar table is generated for the *MAC-MAC* pair. The process leading to the fragmented distribution of scores for 500 games of *MAC* versus *MAC* (displayed in Figure 5), is well depicted in Figure 10. Figure 10 shows a continuation of the new tendency observed in Figure 9; namely, a transition to increasingly unstable event matrices. Recall that the most probable event matrix for the *MAC-MAC* pair is $\{81,9,9,1\}$. This set of values evidently lies in a highly unstable region of the $\{W,X,Y,Z\}$ spectrum, in which probabilistic fluctuation gives rise to one of three situations. Together, the three situations account for the fragmentation of the *MAC-MAC* pair's distribution of scores.

First, perpetual mutual co-operation can be attained very rapidly, as on move 115 in the $\{79,X,Y,1\}$ block, or even immediately, as on move 101 in the $\{83,X,Y,8\}$ block. The onset of rapid perpetual mutual co-operation, when it occurs, is hastened as the sum W plus Z increases. And when it does occur, it results in very high (though not necessarily equal) scores for both twins, in the 2960–2992 point range. This situation contributes to the prominence at the high end of the distribution in Figure 5.

Second, the onset of perpetual mutual co-operation can be noticeably retarded, occurring anywhere between move 364 and move 396 in Figure 10. The delay increases with the sum of W plus Z . And the delay, when it occurs, marks a disparity in the final scores. One pair-member attains roughly 2800–2950 points; the other, roughly 2200–2500 points. This situation thus contributes to the high-range prominence, and it forms the prominence in the next-lowest point range in Figure 5. The trough from 2600–2800 points occurs, self-evidently, because no probabilistic event matrix in this region of the $\{W,X,Y,Z\}$ spectrum can give rise to a deterministic score in that range.

Third, there may be no onset of perpetual mutual co-operation. Such cases give rise to disparate, low final scores. The range of the disparity varies roughly from 250 points to 550 points. This range increases, between blocks, with the sum W plus Z ; and it increases, within blocks, with the difference X minus Y . A typical score is 1621–1306 points. This

Initial W,X,Y,Z	Perpetual (C,c)	Final Score	Initial W,X,Y,Z	Perpetual (C,c)	Final Score	Initial W,X,Y,Z	Perpetual (C,c)	Final Score
79,10,10,1	move 115	2960-2960	81,9,9,1	move 113	2965-2965	83,8,8,1	move 111	2970-2970
79,11,9,1	move 364	2352-2887	81,10,8,1	none	1330-1585	83,9,7,1	none	1318-1623
79,12,8,1	none	1330-1565	81,11,7,1	none	1318-1603	83,10,6,1	none	1306-1641
79,13,7,1	none	1318-1583	81,12,6,1	none	1306-1621	83,11,5,1	move 396	2256-2931
79,14,6,1	none	1306-1601	81,13,5,1	move 385	2278-2933	83,12,4,1	none	1285-1680
79,10,9,2	move 364	2352-2960	81,9,8,2	none	1330-1585	83,8,7,2	none	1318-1623
79,11,8,2	none	1330-1565	81,10,7,2	none	1318-1603	83,9,6,2	none	1306-1641
79,12,7,2	none	1318-1583	81,11,6,2	none	1306-1621	83,10,5,2	move 396	2256-2931
79,13,6,2	none	1306-1601	81,12,5,2	move 385	2278-2933	83,11,4,2	none	1285-1680
79,14,5,2	move 375	2298-2933	81,13,4,2	none	1285-1660	83,12,3,2	none	1272-1702
79,9,9,3	move 111	2965-2965	81,8,8,3	move 109	2970-2970	83,7,7,3	move 107	2975-2975
79,10,8,3	none	1330-1565	81,9,7,3	none	1318-1603	83,8,6,3	none	1306-1641
79,11,7,3	none	1318-1583	81,10,6,3	none	1306-1621	83,9,5,3	move 396	2256-2931
79,12,6,3	none	1306-1601	81,11,5,3	move 385	2278-2933	83,10,4,3	none	1285-1680
79,13,5,3	move 375	2298-2933	81,12,4,3	none	1285-1660	83,11,3,3	none	1272-1702
79,8,8,5	move 107	2970-2970	81,7,7,5	move 105	2975-2975	83,6,6,5	move 104	2978-2978
79,9,7,5	none	1318-1583	81,8,6,5	none	1306-1621	83,7,5,5	move 396	2256-2931
79,10,6,5	none	1306-1601	81,9,5,5	move 385	2278-2933	83,8,4,5	none	1285-1680
79,11,5,5	move 375	2298-2933	81,10,4,5	none	1285-1660	83,9,3,5	none	1272-1702
79,12,4,5	none	1285-1640	81,11,3,5	none	1272-1682	83,10,2,5	none	1259-1724
79,7,6,8	none	1306-1601	81,6,5,8	move 101	2976-2981	83,5,4,8	move 101	2977-2982
79,8,5,8	move 375	2298-2933	81,7,4,8	none	1285-1660	83,6,3,8	move 101	2972-2987
79,9,4,8	none	1285-1640	81,8,3,8	none	1272-1682	83,7,2,8	move 101	2967-2992
9,10,3,8	none	1272-1662	81,9,2,8	none	1259-1704	83,8,1,8	none	1245-1750
79,11,2,8	none	1259-1684	81,10,1,8	none	1245-1730	83,9,0,8	none	1231-1776

Figure 10: MAC versus MAC, varying event matrices and scores.

situation contributes to the two other prominences, in the 1500–1700 and 1300 point ranges of Figure 5. Again, troughs occur in the 1900–2200 and 1000–1200 point ranges because such scores are deterministically inaccessible from the event matrices in this probabilistic region of the $\{W,X,Y,Z\}$ spectrum.

These three different situations occur consecutively in the $\{83,X,Y,5\}$ block of Figure 10. The instability of the event matrix is well evidenced in this block. The matrix $\{83,6,6,5\}$ gives rise to perpetual mutual co-operation on move 104, and results in a final score tied at 2978 points. When the (X,Y) values fluctuate from $(6,6)$ to $(7,5)$, perpetual mutual co-operation does not begin until move 396, with a resultant score of 2256–2931. One further fluctuation in (X,Y) values, from $(7,5)$ to $(8,4)$, debars further perpetual mutual co-operation from occurring in this block, and results in scores such as 1285–1680. Thus, in this block, an initial (X,Y) difference of only 4 causes severe decrements, of 1693 and 1298 points, to the final scores of the *MAC-MAC* pair.

In sum, Figures 8, 9, and 10 account for the different non-normal distributions of final scores in repeated encounters between *MEU-MEU*, *MAE-MAE* and *MAC-MAC* pairs. Moreover, these tables reveal some unexpected, interesting and shifting tendencies across the spectrum of possible event matrices. These tendencies convey an appreciation of the general nature of the relationship between the probabilistic and deterministic phases of the maximization family's play.

This appreciation extends to cases in which siblings, rather than twins, are paired. One need not resort to further analyses of numerous representative probabilistic fluctuations, but one might outline just one case to illustrate how the understanding can be applied. One hundred games of *MAC* versus *MAE* generate the non-normal distributions of final scores displayed in Figure 11.

The most probable event matrix for *MAC* versus *MAE* is $\{64,26,7,3\}$, which gives rise to perpetual mutual co-operation on move 295, and thence to the most probable score of *MAC* 2473, *MAE* 2913. But the average score for 100 games is found to be *MAC* 1849, *MAE* 2123. Again, the distributions explain the discrepancy. But what gives rise to the distributions?

In the initial event matrix, let W and Z be held constant at their most probable respective values of 64 and 3, and let (X,Y) fluctuate from $(25,8)$ to $(29,4)$. The results are displayed in Figure 12, which illustrates how the distributions in Figure 11 arise. The probabilistic event matrices for *MAC* versus *MAE* lie in an unstable region of the $\{W,X,Y,Z\}$ spectrum, from which two main deterministic states are accessible. Perpetual mutual co-operation either commences around move 300, or it does not commence at all. The first state contributes to the higher

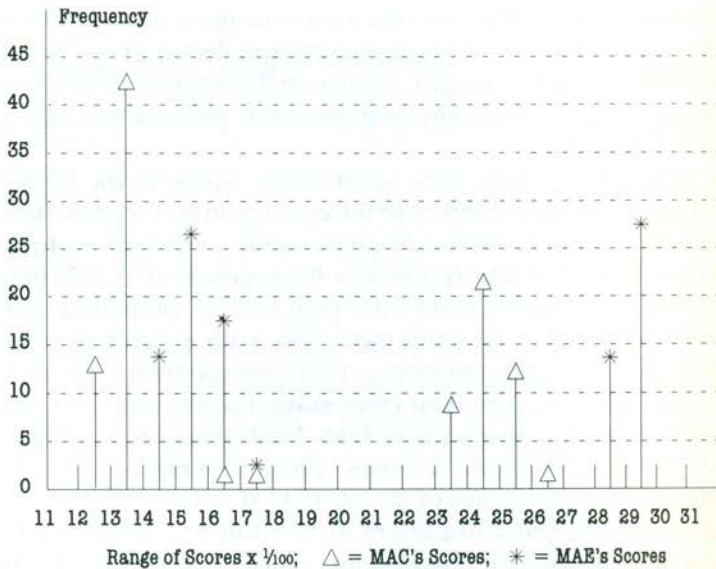


Figure 11: MAE versus MAC. Histogram of scores for 100 games. Average score: MAE 2123, MAC 1849.

point-range features in the respective distributions; the second, to the lower. In the first situation, MAE outpoints MAC by a typical score of 2900–2500; in the second situation, by a typical score of 1600–1350.

6. An Appeal to Evolution

Similar outlines could naturally be drawn to account for the results of other encounters between maximization family siblings. But the foregoing analyses explain the reasons for MAC's relatively poor performances against its twin and its siblings, as revealed in Figure 2. MAC's initially high co-operative weighting, which stands MAC in better stead than its siblings in competition against other strategic families, militates against MAC in intrafamilial competition. MAC's probabilis-

Initial W, X, Y, Z	Perpetual (C, c)	Final Score
64, 25, 8, 3	none	1320 - 1425
64, 26, 7, 3	move 295	2473 - 2913
64, 27, 6, 3	move 299	2457 - 2922
64, 28, 5, 3	move 302	2443 - 2933
64, 29, 4, 3	none	1280 - 1495

Figure 12: MAC versus MAE, varying event matrices and scores.

tic event matrices span an unstable region of the $\{W,X,Y,Z\}$ spectrum, and the instability causes moderate to extreme discrepancies between *MAC*'s most probable and average intrafamilial scores.

MAC's less co-operatively weighted siblings, *MAE* and *MEU*, are also afflicted by this familial syndrome, but to correspondingly lesser extents. *MAD* is immune to it; hence *MAD*'s most probable and average scores coincide. But *MAD*'s immunity is conferred by a property which entails far worse consequences in the tournament environment; namely, the inability to cross the threshold of perpetual mutual co-operation. Hence, *MAD*'s prophylactic measure is more debilitating than the syndrome which it prevents.

Does the lack of sibling recognition among maximization family members lend itself to any social or biological interpretation? One might be tempted to draw a superficial moral from this story, to the effect that since maximization strategies embody the property of exploitiveness, then even if they find no exploitable strategies in their environment they cannot refrain from exploiting one another. Simplistic sociopolitical and ethological allegories abound. One might envision a proverbial pack of thieves falling out over their spoils, instantiating Hobbes's (1651, ch. 13) notion of fleeting or insincere alliances in his natural war "of every man against every man." One might imagine a school of sharks devouring one another during a feeding frenzy, in the spirit of Spencer's (1898, pp. 530-31) "survival of the fittest," which naturally applies to predators as well as to prey.⁴

But these allegorical interpretations do not account for the mathematical niceties of the maximization family's interactions. If exploitiveness were a pivotal determinant of strategic robustness, then *MAD* would out-perform its siblings in intrafamilial competition, followed by *MEU*, *MAE*, and *MAC*. As Figure 2 shows, this does not occur. Moreover, as Axelrod (1980b) predicts and my (1992) tournament demonstrates, the most robust strategy (*MAC*) is able both to exploit the exploitable and to co-operate with the provokable. Hence, if a game-theoretic analogue of sociobiological fitness is strategic robustness, then exploitiveness alone does not make a strategy relatively robust.

A deeper interpretation of the maximization family's performance does not preclude biological and ethological analogies; rather, it suggests that a fundamental comparison be made between species recognition and strategic identification. Mechanisms of species recognition are as yet relatively little-understood across the broad zoological spectrum; however, it appears that many forms of conspecific recognition and subsequent behaviour are mediated by pheromones (e.g., see Stoddart 1976; Birch and Haynes 1982). Hosts of intraspecific biochemical messages are transmitted and received in the animal kingdom - humans included -

and it is easy to appreciate why natural selection would have favoured the evolution of this general mechanism across diverse ranges of species.

While *n*-pair, repeated Prisoner's Dilemma tournaments are susceptible to ecological modeling (see Axelrod 1980b; Marinoff 1992), they are also amenable to evolutionary change. We are not now referring to Maynard-Smith's (1982) evolutionary games theory, which ingeniously models population genetics using game-theoretic constructs;⁵ rather, we are invoking a cognitive scientific approach, which effects strategic evolution by simulating aspects of the neo-Darwinian paradigm using computer technology and high-level programming languages. For example, Koza's (1991) LISP programs generate the co-evolution of minimax strategies in a generic two-person, zero-sum game. Fujiki and Dickinson (1987) adapt genetic algorithms to manipulate LISP expressions, thereby evolving strategies for the Prisoner's Dilemma. Danielson (1992, pp. 133-42) uses PROLOG to simulate both strategic adaptation and learning in the Prisoner's Dilemma. Although evolved strategies may incorporate meta-strategic properties, their paradigmatic development is distinctly evolutionary (rather than meta-game-theoretic⁶). Danielson (1992, pp. 51-52) calls this approach "moral engineering."

None the less, it can be predicted that Axelrod's (1984, p. 15) maxim for the iterated Prisoner's Dilemma, "... there is no best strategy independent of the strategy used by the other player," is unlikely to be threatened by the emergence of an evolutionary "super-strategy." To see why the maxim holds in evolutionary scenarios, consider the following argument. Hypothesize that the maximization family evolved some reliable mechanism of familial identification.⁷ A suitable strategic analogue of a pheromone could be a designated substring of co-operations and defections (e.g., CCCCCDDDDD) nested somewhere within the 100 random moves. When two such evolved maximization siblings compete, they construct their event matrices as usual. But at the same time, they also monitor their opponent's string of moves. When either maximization strategy detects the predetermined substring that identifies its opponent as "conspecific," it immediately sends back the same substring in reply, then initiates perpetual co-operation. Its opponent detects this identifying substring, and responds with perpetual co-operation.

If the maximization strategies had been vouchsafed the capacity for such behaviour, then their performances would have been considerably enhanced. MAE would have finished second instead of third in overall robustness, while MAC's margin of victory would have been even wider. However, this mechanism of identification would not render MAC the "best" strategy, independent of environment. The mechanism could also backfire, in at least three ways.

First, any randomizing strategy (or member of the probabilistic family) could fortuitously generate the predetermined recognition-string, and would thereby elicit perpetual co-operation from a maximizer. The maximization strategy would then be exploited. Second, given the knowledge that the maximization family employs a predetermined recognition-string, and given also a sufficient number and length of encounters, then presumably any evolving strategy could, by trial-and-error, eventually learn to generate the recognition-string itself. Again, the maximization strategy would be exploited. Third, given an evolutionary scenario, one might witness the emergence of a "rogue" maximization strategy, which first produces the identification necessary to elicit perpetual co-operation from its maximizing sibling, and then proceeds to defect perpetually itself, thus maximizing its own long-term gains (except against its twin). Such a strategy, ironically, would be exploiting the very mechanism that evolved to circumvent intraspecific exploitation.

As Danielson (1992, p. 135) notes, "Flexibility makes new predatory tricks possible and requires co-operative players to be more cautious." So, notwithstanding emergent evolutionary models, one confidently predicts the reassertion of the problematic nature of the Prisoner's Dilemma. Ever-more successful strategies may evolve, but – shades of the antlers of the Irish elk – any attribute that guarantees today's success may also seal tomorrow's doom.

Notes

- 1 The intrafamilial data in this table is based on 100 games between siblings, and 500 games between twins.
- 2 As previously noted (Marinoff 1992, p. 214), an occurrence of mutual defection lowers the expected utility of further defection. Thus, for maximization family encounters, mutual defection increases the propensity for mutual co-operation.
- 3 But a tied final score does not result uniquely from this process. For example, the {20,33,32,15} matrix in Figure 8 generates a single unilateral co-operative play (*D,c*) at move 911, which produces a symmetric matrix and hence a tied final score in the absence of mutual co-operation.
- 4 A deeply entrenched fallacy attributes the phrase "survival of the fittest" to Darwin. In fact, it was coined by Spencer in 1863/4 as a synonym for "natural selection." Darwin long resisted adopting this term, despite Wallace's (1866) and others' promptings.
- 5 Contrary to Axelrod's and Hamilton's (1981) assertion, there is no evolutionarily stable strategy in the repeated Prisoner's Dilemma. See Boyd and Lorberbaum (1987), Axelrod and Dion (1988), Marinoff (1990).
- 6 Meta-game theory was formalized by N. Howard (1971).

- 7 J. Howard (1988) lists the source code, in BASIC, of a self-recognizing Prisoner's Dilemma strategy.

References

- Axelrod, R. (1980a). Effective choice in the Prisoner's Dilemma. *Journal of Conflict Resolution*, 24: 3–25.
- (1980b). More effective choice in the Prisoner's Dilemma. *Journal of Conflict Resolution*, 24: 379–403.
- (1984). *The Evolution of Cooperation*. New York: Basic Books.
- Axelrod, R., and D. Dion (1988). The further evolution of cooperation. *Science*, 242: 1385–90.
- Birch, M., and K. Haynes (1982). *Insect Pheromones*. The Institute of Biology's Studies in Biology, no. 147. London: Edward Arnold.
- Boyd, R., and J. Lorberbaum (1987). No pure strategy is evolutionarily stable in the repeated Prisoner's Dilemma game. *Nature*, 327: 58–59.
- Danielson, Peter (1992). *Artificial Morality*. London: Routledge.
- Fujiki, C., and J. Dickinson (1987). Using the genetic algorithm to generate lisp source code to solve the Prisoner's Dilemma. In J. Grefenstette (ed.), *Genetic Algorithms and Their Applications: Proceedings of the Second International Conference On Genetic Algorithms* (Hilldale, NJ: Lawrence Erlbaum), pp. 236–40.
- Hobbes, T. (1957 [1651]). *Leviathan*. Edited by M. Oakeshott. Oxford: Basil Blackwell.
- Howard, J. (1988). Cooperation in the Prisoner's Dilemma. *Theory and Decision*, 24: 203–13.
- Howard, N. (1971). *Paradoxes of Rationality: Theory of Metagames and Political Behaviour*. Cambridge, MA: MIT Press.
- Koza, J. (1991). Evolution and co-evolution of computer programs to control independently-acting agents. In J. Meyer and S. Wilson (eds.), *From Animals to Animats: Proceedings of the First International Conference on Simulation of Adaptive Behaviour* (Cambridge: MIT Press), pp. 366–75.
- Marinoff, L. (1990). The inapplicability of evolutionarily stable strategy to the Prisoner's Dilemma. *The British Journal for the Philosophy of Science*, 41: 462–71.
- (1992). Maximizing expected utilities in the Prisoner's Dilemma. *Journal of Conflict Resolution*, 36: 183–216.
- Maynard Smith, John (1982). *Evolution and the Theory of Games*. Cambridge: Cambridge University Press.
- Spencer, H. (1898 [1863–4]). *The Principles of Biology*, Vol. 1. London: Williams and Norgate.
- Stoddart, D. (1976). *Mammalian Odours and Pheromones*. The Institute of Biology's Studies in Biology, no. 173. London: Edward Arnold.