Maximizing Expected Utilities in the Prisoner's Dilemma

LOUIS MARINOFF
Centre for Applied Ethics, University of British Columbia

This article reports the results of a computer experiment with iterated prisoner's dilemmas conducted as an interactive tournament of competing strategies and families of strategies. The purposes of the experiment are to complement Axelrod's previous tournaments and to supplement his findings. For his competitions, Axelrod drew on an unregulated population of strategies. In contrast, the interactive tournament regulates the composition of the strategic population itself. By grouping the competing strategies into families, whose members are related in certain ways, the performance characteristics of particular strategies are studied by varying parameters in their familial program logic. By this means, optimal strategic performance can be "bred" into domesticated populations. Two new methods are developed for assessing strategic robustness: combinatorial analysis and eliminatory ecosystemic competition. The strategy that maximizes expected utility with the most cooperative initial weighting is found to be most robust in the interactive environment.

Axelrod's (1980a, 1980b) computer tournaments for the iterated prisoner's dilemma revealed salient criteria of effective (and ineffective) strategic performance in the two-player, multiple-pair context. In order to amplify Axelrod's findings and to test his principal conclusions, I conducted a further computer experiment called the "interactive tournament."

In his previous tournaments, Axelrod regulated three of four key environmental factors: the payoff structure, the number of iterations in a game, and the players' knowledge of this number. He then solicited unregulated populations of competing strategies from diverse sources. His strategic populations were thus of the "wild" type.

The alternative to a wild population is, of course, a "domesticated" one. If the strategic types in competition are themselves regulated, the experimenter then exercises fuller control over the tournament environment. Domesticated strategies can be "bred" that incorporate or exclude virtually any

JOURNAL OF CONFLICT RESOLUTION, Vol. 36 No. 1, March 1992 183-216

TABLE 1
Prisoner's Dilemma Payoff Matrix

		Column Player	
		c	*d*
Row Player	*C*	3,3	0,5
	D	5,0	1,1

NOTE: *C* and *c* denote "cooperation"; *D* and *d* denote "defection."

combination of niceness, provocability, forgiveness, and exploitiveness, among other qualities. And certain wild strategies, selected for robustness or for any desired performance characteristic, can be maintained "in captivity" and induced to compete against the domesticated strategies. By studying the interactions of captive and domesticated strategies under varying conditions, one can identify properties that tend to make a given strategy more (or less) effective in a particular strategic population within a defined environment. A flexible strategy can be modified until its performance is optimally effective in the context of its competitors and surroundings.

The key environmental factors in the interactive tournament are defined as follows:

1. The payoff structure is identical to that employed by Axelrod in his previous tournaments.
2. The number of iterations per game is held constant at 1,000. This increase (over Axelrod's constant 200 in his first tournament and probabilistically averaged 200 in his second tournament) grants slowly developing strategies an opportunity to attain optimal-performance levels.
3. All strategies are programmed with the property of "integrity"; that is, each strategy adheres to its normal decision rule for the full length of the game. No strategy deviates from its decision rule by making unprovoked late-game defections hoping to exploit both intrinsically cooperative strategies and strategies too slow to retaliate.
4. The 20 strategies competing in the interactive tournament are grouped into "families" in which members are related either closely (by program structure), or more distantly (by conceptual function).

THE STRATEGIC FAMILIES

The five families in the interactive tournament and their members' acronyms and decision rules are summarized as follows:

THE PROBABILISTIC FAMILY

Members of this family cooperate and defect randomly according to their individual probabilistic weightings. The two pure strategies (pure cooperation and pure defection) are included in this family because their program structure is identical to that of the other members. The members' decision rules thus differ by a sole parameter: the probability of cooperation on a given move. This is the only family in the tournament whose members make their moves without taking their opponent's moves into account.

DDD: This is the strategy of pure defection. On every move, DDD cooperates with a probability of zero and defects with a probability of unity.

TQD: This is the strategy of three-quarter random defection. On every move, TQD cooperates with a probability of 1/4 and defects with a probability of 3/4.

RAN: This is the strategy of random equiprobability. On every move, RAN cooperates or defects with a probability of 1/2.

TQC: This is the strategy of three-quarter random cooperation. On every move, TQC cooperates with a probability of 3/4 and defects with a probability of 1/4.

CCC: This is the strategy of pure cooperation. On every move, CCC cooperates with a probability of unity and defects with a probability of zero.

THE TIT-FOR-TAT FAMILY

Members of this family are all related to tit-for-tat, and hence share a similar program structure. Small variations in members' decision rules can naturally result in large variations in competitive performance.

TFT: Tit-for-tat is the primogenitor of the family, and was the most robust strategy in Axelrod's tournaments. TFT cooperates on the first move and plays next whatever its opponent played previously.

TTT: Tit-for-two-tats is less provocable than TFT. TTT would have won Axelrod's first tournament (had it competed), but fared less well in the second. TTT cooperates on the first two moves and defects only after two consecutive defections by its opponent.

BBE: This strategy attempts to "burn both ends" of the strategic candle. It plays exactly as TFT with one modification: BBE responds to an opponent's cooperative move by cooperating with a probability of 9/10. BBE thus attempts to out-perform TFT by being equally provocable but less reliably forgiving.

SHU: This is Shubik's strategy, which ranked fifth in Axelrod's first tournament. It plays as TFT with a modification. SHU defects once following an opponent's first defection, then cooperates. If the opponent defects on a second occasion when SHU cooperates, SHU then defects twice before resuming cooperation. After each occasion on which the opponent defects when SHU cooperates, SHU increments its retaliatory defections by one. SHU thus becomes progressively less forgiving in direct arithmetic relation to the number of occasions on which SHU's cooperation meets with an opponent's defection.

TABLE 2
Event Matrix for Maximization Strategy versus Opponent

		Opponent	
		c	*d*
Maximization strategy	*C*	*W*	*X*
	D	*Y*	*Z*

NOTE: W = number of occasions on which outcome (C, c) obtained
X = number of occasions on which outcome (C, d) obtained
Y = number of occasions on which outcome (D, c) obtained
Z = number of occasions on which outcome (D, d) obtained

TAT: Tat-for-tit is the binary complement of TFT. TAT defects on its first move, then plays next the opposite of whatever its opponent played previously. TAT thus defects in response to cooperation and cooperates in response to defection. TAT has been bred to exhibit contrariness.

THE MAXIMIZING FAMILY

All members of this family maximize expected utilities, but do so with different initial probabilistic weightings. Each member plays randomly for 100 moves (cooperating or defecting according to its particular weighting), and keeps track of all moves made by both itself and its opponent. After 100 moves, an "event matrix" of joint outcome frequencies is used to assign *a posteriori* probabilisties in the calculation of expected utilities for the 101st move and all moves thereafter. The generalized event matrix takes the form shown in Table 2.

From move 101 onward, the maximization strategy computes the probability that each game state obtains as an outcome frequency from the event matrix. By summing the products of the probability of each game state and the payoff of that state, the expected utilities of cooperation and defection are evaluated explicitly as:

$$EU(C) = 3W/(W + X)$$
$$EU(D) = (5Y + Z)/(Y + Z).$$

If $EU(C)$ is greater than or equal to $EU(D)$, the maximization strategy cooperates on move 101; otherwise, it defects. The maximization strategy continues to record outcomes throughout the game, and thus updates the event matrix after every outcome.

The program structure is identical for every member of this family. The critical parameter, in whose value the members differ, is the weight accorded to the probability of a member's random cooperation during the first 100 moves. The maximization family was bred to represent a range of weights.

MEU: Maximization of expected utility is the familial prototype. A reactive-maximization strategy appeared in Axelrod's tournaments under the name of its submitter, DOWNING. DOWNING ranked 10th among 15 entries in the first tournament and 40th among 63 in the second. DOWNING adopted the principle of insufficient reason[1] and assigned (*a priori*) equiprobabilities of 1/2 to each opponent's choice prior to its first move. It then updated the probabilities according to the relative frequencies of actual outcomes. DOWNING would have finished first (in the first tournament) had its initial probabilistic outlook been more optimistic. MEU randomly cooperates or defects with probability 1/2 for the first 100 moves. But in contrast to DOWNING, MEU assumes nothing about the play of its opponent. Instead, MEU notes its opponent's choice on each move and records each joint outcome in the event matrix.

MAD: This strategy maximizes expected utilities with initial weighting at defection. MAD plays exactly as MEU except that, on each of its first 100 moves, MAD defects with a probability of 9/10 and cooperates with a probability of 1/10.

MAE: This strategy maximizes expected utilities with initial weighting at equalized expectation. The numerical values of this weighting are thus dependent on the payoff structure of the game. Given the payoffs of the interactive tournament, MAE cooperates with a probability of 5/7 and defects with a probability of 2/7.

MAC: This strategy maximizes expected utilities with initial weighting at cooperation. MAC plays exactly as MEU except that, on each of its first 100 moves, MAC cooperates with a probability of 9/10 and defects with a probability of 1/0.

THE OPTIMIZATION FAMILY

Unlike the preceding strategic families, members of the optimization family are related neither by common program structures nor by variations on a common decision rule. The attribute shared by this family's members is their demonstrated success in previous competition(s), achieved by implementing decision rules that attempt to optimize future outcomes in light of past ones.

NYD: This is Nydegger's strategy. It ranked third in Axelrod's first tournament. NYD is succinctly described by Axelrod [1980a, 22]:

1. "Alternatives are to be judged equiprobable if we have no reason to expect or prefer one over the other" (Weatherford 1982, 29; see also Luce and Raiffa 1957, 284).

The program begins with tit for tat for the first three moves, except that if it was the only one to cooperate on the first move and the only one to defect on the second move, it defects on the third move. After the third move, its choice is determined from the 3 preceding outcomes in the following manner. Let *A* be the sum formed by counting the other's defection as 2 points and one's own as 1 point, and giving weights of 16, 4 and 1 to the preceding three moves in chronological order. The choice can be described as defecting only when A equals 1, 6, 7, 17, 22, 23, 26, 29, 30, 31, 33, 38, 39, 45, 49, 54, 55, 58, or 61.

GRO: This is Grofman's strategy. It ranked fourth in Axelrod's first tournament. GRO cooperates on the first move. After that, GRO cooperates with probability 2/7 following a dissimilar joint outcome (either [*C*, *d*] or [*D*, *c*]), and always cooperates following a similar joint outcome (either [*C*, *c*] or [*D*, *d*]).

CHA: This is Champion's strategy. It ranked second in Axelrod's second tournament. CHA cooperates on the first 10 moves and plays tit-for-tat on the next 15 moves. From move 26 onward, CHA cooperates unless all of the following conditions are true: The opponent defected on the previous move, the opponent's frequency of cooperation is less than 60%, and the random number between zero and one is greater than the opponent's frequency of cooperation.

ETH: This is Eatherly's strategy. It ranked fourteenth in Axelrod's second tournament, but proved quite robust in a tournament conducted privately by Eatherly himself.[2] ETH cooperates on the first move and keeps a record of its opponent's moves. If its opponent defects, ETH then defects with a probability equal to the relative frequency of the opponent's defections.

THE HYBRID FAMILY

The members of this family share the common attribute that their decision rules, as implied by the family name, are formed by the hybridization of other strategic pairs. This family consists of one "pure" hybrid (bred from two pure strategies) and one "mixed" hybrid (bred from two mixed strategies).

FRI: This is Friedman's strategy, which ranked seventh in Axelrod's first tournament. FRI cooperates until its opponent defects, after which FRI defects for the rest of the game. Hence FRI is both nice and provocable, but completely unforgiving. Its properties in other contexts are elsewhere discussed.[3] FRI's sequence of choices consists either in a string that is identical to CCC, or else in a string that is identical to CCC up to some move, and identical to DDD thereafter. Thus FRI is a pure strategic hybrid.

TES: This is a strategy called "Tester," submitted by Gladstein. TES finished only 46th in Axelrod's second tournament, but proved adept at exploiting potentially successful strategies, thus compromising their would be robustness. TES defects on the first move. If its opponent ever defects, TES "apologizes" by cooperating and plays tit-for-tat thereafter. Until its opponent defects, TES

2. Cited by Axelrod (1980b).

3. E.g. see Harris (1969), Friedman (1971).

> defects with the maximum possible relative frequency that is less than 1/2, not counting its first defection. TES appears somewhat "opportunistic" in character. On the one hand, it attempts to exploit cooperative strategies without being excessively provocative. On the other, it attempts to appease provocable strategies, while retaining its capacity to retaliate. In sum, TES incorporates two mixed strategies: defection with relative frequency up to one-half and TFT. Thus TES is a mixed strategic hybrid.

This completes a description of the 20 competing strategies in the interactive tournament and their classification by common characteristics. It should be stressed that the familial organization employed herein is quite heuristic; any such collection of strategies can be grouped in a large number of ways.

For example, one might choose niceness (the property of never being the first to defect) as a criterion of distinction. CCC, TFT, TTT, SHU, NYD, GRO, CHA, ETH, and FRI are nice strategies; DDD, TAT, and TES can be termed *rude* strategies (where rudeness is the property of always being the first to defect). This leaves TQD, RAN, TQC, BBE, MEU, MAD, MAE, and MAC unqualified, for these strategies are neither nice nor rude. They might be assigned the predicate of *nide* (where *nideness* is the property of being indeterminate with respect to primacy of defection).

In sum, although the five strategic families do not constitute a rigorous or exhaustive system of classification, they are useful as heuristic aids in a controlled experiment. A tournament whose population is of the wild variety has no express need of such groupings; as in Axelrod's experiments, the idea is to observe the competition of an unregulated population and to see which strategies are successful in a free-for-all environment. In a tournament whose population is of the domesticated and captive varieties, however, these familial groupings allow the observation of the relative success of various strategic shadings, whether across the spectrum of a single parameter in a common program structure, or in terms of conceivable variations on a common functional theme.

COMBINATORIAL ROBUSTNESS

The strategies played round-robin against one another and their twins. The matrix of raw scores and the standings to which they give rise are shown in Tables 3 and 4 respectively.[4]

4. The raw scores are based on single games of 1,000 moves except in the case of the maximization family, whose intrafamilial scores were found to be abnormally distributed. Scores

TABLE 3
Matrix of Raw Scores, Main Tournament

	DDD	TQD	RAN	TQC	CCC	TFT	TTT	BBE	SHU	TAT
DDD	1000	1952	2992	3996	5000	1004	1008	1004	1176	4996
TQD	762	1727	2634	3550	4470	1673	2324	1520	948	3580
RAN	502	1354	2243	3139	3972	2193	3129	2098	713	2299
TQC	251	1095	1914	2685	3472	2706	3295	2436	537	1124
CCC	0	795	1542	2292	3000	3000	3000	2700	3000	0
TFT	999	1673	2193	2701	3000	3000	3000	1036	3000	2250
TTT	998	1444	1874	2365	3000	3000	3000	2662	3000	1800
BBE	999	1690	2433	2766	3200	1041	3197	1033	1174	2367
SHU	956	1878	2913	3877	3000	3000	3000	974	3000	4529
TAT	1	1055	2219	3534	5000	2250	2800	2132	384	2000
MEU	947	1995	2899	3940	4920	1113	1538	1133	1180	4750
MAD	994	2037	3004	3912	4980	1024	1147	1013	1168	4972
MAE	932	1970	2903	3899	4838	1272	1624	2544	1183	4628
MAC	905	1878	2900	3750	4824	2965	2995	2665	1237	4556
NYD	334	764	1543	2305	3000	3000	3000	2637	3000	14
GRO	427	1233	2111	2721	3000	3000	3000	2385	3000	1156
CHA	990	1541	2010	2281	3000	3000	3000	2660	3000	1697
ETH	999	1475	1875	2445	3000	3000	3000	2640	3000	1681
FRI	999	2023	2933	4033	3000	3000	3000	1027	3000	4991
TES	999	1688	2156	2731	4000	2999	4000	1044	2999	2246

The winner of the main tournament, by a comfortable margin, is MAC. MAC is the most cooperatively weighted member of the maximization family. Second place, by a narrower margin, goes to MAC's closest relative, MAE. Third place is taken by the least-forgiving member of the tit-for-tat family, SHU. Fourth place belongs to the pure hybrid, FRI. Fifth and sixth places are occupied by members of the optimization family, CHA and ETH.

The average scores (per game) are distributed within the following limits. Because the length of a game is 1,000 moves, the maximum achievable score in any game is 5,000 points; the minimum, 0 points. These extrema occur if one strategy defects 1,000 times, and its opponent cooperates 1,000 times. This extreme situation actually obtained in two cases: DDD and TAT both scored maximum points against CCC, which went scoreless against both. While these two dismal outings by CCC contributed to its last-place finish,

between maximization siblings are averaged over 100 games; between maximization twins, over 500 games.

MEU	*MAD*	*MAE*	*MAC*	*NYD*	*GRO*	*CHA*	*ETH*	*FRI*	*TES*
1212	1024	1272	1380	3664	3292	1040	1004	1004	1004
920	777	1025	1098	4484	3023	2426	2405	748	1693
659	529	743	825	3968	2681	3200	3076	523	2161
405	307	479	570	3460	2586	3476	3280	248	2721
120	30	243	264	3000	3000	3000	3000	3000	1500
1108	1019	1267	2965	3000	3000	3000	3000	3000	2999
1143	1002	1204	2935	3000	3000	3000	3000	3000	1500
1148	1018	2989	3140	3242	2735	3215	3225	1027	1049
1125	1008	1263	1322	3000	3000	3000	3000	3000	2999
230	42	368	466	4984	3266	2837	2851	6	2251
2384	1003	2396	1887	4875	3241	2610	2294	955	1175
1181	1029	1266	1332	4940	3273	1193	1243	996	1013
2356	987	2594	2123	4852	3232	2870	3000	939	1312
1741	971	1849	1807	4814	3270	3000	2893	955	2926
165	45	217	279	3000	3000	3000	3000	3000	2500
586	468	672	670	3000	3000	3000	3000	3000	2995
2140	988	2215	2210	3000	3000	3000	3000	3000	2987
1614	988	2263	2505	3000	3000	3000	3000	3000	2999
1195	1031	1259	1325	3000	3000	3000	3000	3000	1007
1175	1008	1307	2951	2500	2995	3007	2999	1002	2998

neither of the two strategies that exploited CCC to the limit fared much better than their victim overall.

A useful bench mark is the 3,000 point level, attained by both members of any strategic pair that practices mutual cooperation for an entire game. This occurred on 81 occasions in all possible encounters between nice strategies (CCC, TFT, TTT, SHU, NYD, GRO, CHA, ETH, and FRI). Owing to the mixture of nice, rude, and nide strategies in the population, no strategy – nice or otherwise – was able to maintain an average score of 3,000 points. MAC and MAE, which fared best with respective averages of 2,645 and 2,503 points per game, are neither nice nor rude, but nide. SHU, the best of the nice strategies, managed an average of 2,492.

With respect to points allowed, CCC, TQC, NYD, and GRO surpassed the 3,000 point bench mark. On this side of the ledger, the accomplishment is of dubious merit. It indicates that these four strategies are the most exploitable.

TABLE 4
Main Tournament, Ranks and Scores

Rank (Offense)	*Strategy*	*Points Scored*	*Average Score*	*Points Allowed*	*Average Allowed*	*Rank (Defense)*
1	MAC	52901	2645	32054	1603	6
2	MAE	50058	2503	26891	1345	4
3	SHU	49844	2492	39699	1985	9
4	FRI	48823	2441	35403	1770	7
5	CHA	48719	2436	55874	2794	16
6	ETH	48484	2424	55270	2764	14
7	MEU	47235	2362	22607	1130	3
8	TFT	47210	2361	47240	2362	11
9	TES	46804	2340	41789	2089	10
10	TTT	45927	2296	54057	2703	13
11	BBE	42688	2134	37343	1867	8
12	GRO	42424	2121	60594	3030	17
13	TQD	41787	2089	31267	1563	5
14	MAD	41717	2086	15274	764	2
15	DDD	40024	2001	14994	750	1
16	RAN	40007	2000	47291	2365	12
17	TAT	38676	1934	55636	2782	15
18	NYD	37803	1890	72783	3639	19
19	TQC	37047	1852	62922	3146	18
20	CCC	36486	1824	75676	3784	20

Surprisingly, 4 of the top 10 strategies (CHA, ETH, TFT, and TTT) were out-scored, on average, by their opponents. But crude averages can be misleading. The relative success of these strategies lies in the precise distributions and magnitudes of their individual scores.

MAC is the most successful strategy in the main tournament involving 20 strategies. Next, one must ask: How robust is MAC in the interactive environment?

In his second tournament, Axelrod used an interesting method involving step-wise regression to assess strategic robustness. He found TFT to be the most robust strategy in that environment.

This study adopts a different methodology, that of combinatorial analysis. Given a set of n elements, one can combine r elements from that set in $n!/r!(n-r)!$ distinct ways. In the interactive tournament, the number of strategies (or elements) is 20. The 20 strategies, of course, can be combined in just one way. But r can assume a range of theoretical values, from $1 \leq r \leq n$. In practice, at least two strategies are required for a competition to take place, so the value $r = 1$ is not applicable here.

The total number of possible subtournaments that can be conducted, from all combinations of strategies for each applicable value of r, is 616,666. The total number of subtournaments in which each strategy competes is 524,287. In order to evaluate the results of this large number of subtournaments, the following procedure is adopted. All subtournament combinations involving r strategies are conducted, one at a time, for each value of r.

Let r have a given value. Suppose strategy S_i ranks first in the first subtournament conducted (for that r). Then strategy S_i is awarded $(r - 1)$ points. Similarly, if strategy S_j ranks second, then strategy S_j fared better than $(r - 2)$ other strategies in that particular combination. Hence, strategy S_j is awarded $(r - 2)$ points. This procedure is applied to all strategies in that subtournament combination. In other words, each strategy in that particular combination is awarded a number of points equal to the number of strategies it betters. Suppose strategy S_k ranks last. Because S_k betters no strategies, it is awarded no points.

The second subtournament combination involving r strategies (for the same value of r) is then tried. Once again, points are awarded to each strategy appearing in this combination, according to the number of other strategies it betters, from $(r - 1)$ points for the first-ranking strategy to zero points for the last-ranking strategy.

When a given subtournament combination consists of nice strategies only, they all achieve identical scores. In such cases, when r nice strategies draw, they each receive $(r - 1)$ points. And most generally, if any subtournament involving r strategies sees p of these strategies tied for q^{th} place, then each of the p strategies receives $(r - q)$ points.

After $C(20, r)$ different combinations are exhausted for the given r, each strategy will have appeared in $19!/(r-1)!(20-r)!$ different subtournaments. In order to determine which strategy is most successful for this value of r, the *efficiency* of each strategy's performance is calculated according to the following formula. If a strategy wins each and every subtournament for this value of r, its point awards would total

$$19!/(r-2)!(20-r)!$$

This is the maximum number of points awardable to a strategy for any given value of r. The relative efficiency of a strategy, then, is simply its actual point award total divided by this maximum number. (The relative efficiency is then multiplied by 100 for expression as an efficiency percentage.)

TABLE 5
19 Appearances in 20 Subtournaments Involving 19 Strategies

	Rank									
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
MAC	19	0	0	0	0	0	0	0	0	0
MAE	0	12	3	2	0	2	0	0	0	0
SHU	1	6	9	1	2	0	0	0	0	0
FRI	0	0	6	8	1	3	0	0	1	0
CHA	0	2	0	6	6	4	1	0	0	0
ETH	0	0	1	1	8	4	4	1	0	0
MEU	0	0	1	2	1	5	2	3	1	4
TFT	0	0	0	0	1	2	10	5	1	0
TES	0	0	0	0	1	0	3	8	7	0
TTT	0	0	0	0	0	0	0	3	10	6
BBE	0	0	0	0	0	0	0	0	0	4
GRO	0	0	0	0	0	0	0	0	0	3
MAD	0	0	0	0	0	0	0	0	0	3
TQD	0	0	0	0	0	0	0	0	0	0
RAN	0	0	0	0	0	0	0	0	0	0
DDD	0	0	0	0	0	0	0	0	0	0
TAT	0	0	0	0	0	0	0	0	0	0
NYD	0	0	0	0	0	0	0	0	0	0
TQC	0	0	0	0	0	0	0	0	0	0
CCC	0	0	0	0	0	0	0	0	0	0

A specific example of the entire procedure is illustrated in Table 5 for the 20 different subtournaments conducted by forming all possible combinations of 19 strategies.

For $r = 19$, there are 20 possible subtournaments. Each strategy appears in 19 subtournaments, and can be awarded a maximum of 18 points in each appearance. Hence, the ideal point award total is $19 \times 18 = 342$ total points.

Because MAC ranked first in all its appearances, it actually achieved this ideal; hence, its efficiency is 100% in subtournaments involving 19 strategies.

MAE ranked second in 12 subtournaments; third in three subtournaments; fourth in two subtournaments; sixth in two subtournaments. Hence, MAE bettered 17 opponents on 12 occasions; 16 on three occasions; 15 on two occasions; and 13 on two occasions. This tally accounts for MAE's 19 appearances. MAE's relative efficiency is therefore

$$[(12 \times 17) + (3 \times 16) + (2 \times 15) + (2 \times 13)]/342 = 308/342 = .901$$

				Rank					
11	12	13	14	15	16	17	18	19	E%
0	0	0	0	0	0	0	0	0	100
0	0	0	0	0	0	0	0	0	90.1
0	0	0	0	0	0	0	0	0	89.8
0	0	0	0	0	0	0	0	0	81.6
0	0	0	0	0	0	0	0	0	79.5
0	0	0	0	0	0	0	0	0	74.3
0	0	0	0	0	0	0	0	0	66.7
0	0	0	0	0	0	0	0	0	65.8
0	0	0	0	0	0	0	0	0	60.8
0	0	0	0	0	0	0	0	0	54.7
8	5	2	0	0	0	0	0	0	43.0
5	5	5	1	0	0	0	0	0	40.1
3	4	3	5	0	1	0	0	0	37.4
4	6	9	0	0	0	0	0	0	36.4
0	0	1	7	10	1	0	0	0	24.6
0	0	0	6	10	2	0	0	1	22.2
0	0	0	0	0	16	1	2	0	15.2
0	0	0	1	0	0	16	2	0	11.4
0	0	0	0	0	0	2	15	2	5.6
0	0	0	0	0	0	1	1	17	0.9

Thus MAE is 90.1% efficient in subtournaments involving 19 strategies.

In Table 5, notice that the nonzero entries tend to be clustered along the main diagonal of the matrix. This general lack of dispersion throughout each row indicates that a given strategy tends to achieve the same rank, or else to perform within a narrow range of ranks, in each of its appearances. One extreme case is MAC, which ranked first in the 19 subtournaments in which it appeared. At the other extreme is MEU, whose rankings are distributed across eight consecutive columns. In its 19 appearances, MEU attained a range of ranks between third and tenth places inclusive.

The average rank dispersion in Table 5 (that is, the average number of different ranks attained by a given strategy), is 4.6 of a possible 19 ranks per strategy. Overall, the actual rank attainments are dispersed over less than 25% of the field of possible rank attainments. This denotes an expected result; namely, that in the 20 subtournaments involving different combinations of 19 strategies, the absence of any particular strategy from a given subtournament does not drastically influence the relative success of the remaining competitors. In other words, slight variations in the constitution of a large population do not exert a pronounced effect on the bulk of its members' performances.

By the same token, one observes an increasing dispersion of rankings as the number of strategies per subtournament diminishes and the corresponding number of possible combinations increases. In the distribution of rankings for the 184,756 combinations of 10 strategies (in which each strategy appears in 92,378 subtournaments and is absent from a like number) the average rank dispersion reaches 9.6 of a possible 10 ranks per strategy, and actual rank attainments are dispersed over 96% of the field of possible rank attainments. The large number of combinations of 10 strategies allows great variation in relative performance.

It transpires that MAC dominates all group sizes from 20 down to 7 competitors inclusive. MAE dominates groups of 6 and 5 competitors, and FRI prevails in groups of 4 and 3. In the 190 subtournaments involving 2 strategies, wherein each strategy makes 19 appearances, FRI, SHU and TFT are most efficient.

These results can be summarized as follows. A total of 616,666 different subtournaments have been conducted by taking all combinations of the population of competing strategies in all group sizes from 20 to 2 competitors. In all, each strategy appears in 524,287 subtournaments (the sum of its appearances in each group size), and the efficiency of each strategy's performance was found for each group. A relative measure of robustness can now be made by calculating each strategy's overall efficiency across the entire range of group sizes.

A strategy's overall efficiency is simply the weighted average of its relative efficiencies in all groups. Suppose a given strategy appears in N_i subtournaments for all combinations $C(i, 20)$ of i competitors, and attains a relative efficiency of E_i in that group. Then the given strategy's overall efficiency, E_o is found by

$$E_o = \sum_i E_i \times N_i \div \sum_i N_i$$

(where the denominator = 524,287). The results of this calculation for all strategies appear in Table 6.

Comparing the standings in Table 6 (overall efficiencies) and 4 (main tournament results), it seems significant that the upper six and lower six strategies maintain identical ranks in both cases. Given that Table 4 is the result of the unique subtournament involving the single combination of 20 strategies, and that Figure 1 is the weighted result of 616,666 different subtournaments involving all combinations of all groups, then the upper and lower third of the compiled standings of more than 600,000 subtournaments are "determined," as it were, by the unique outcome featuring the largest group. It is a matter of speculation whether such determination would obtain anew, and to what degree, in different strategic populations.

The combinatorial analysis of subtournaments concludes with a graph (Figure 1) that illustrates how the efficiencies of the upper six strategies change as a function of group size.

MAC and MAE are the sole top strategies whose efficiencies increase uniformly with the size of the competing group. SHU and FRI, which rank third and fourth respectively, do well because their efficiencies increase after falling off sharply in smaller groups. CHA and ETH, which rank fifth and sixth respectively, experience a less sharp early decrease in smaller groups, a gradual increase in mid-sized groups, and a gradual falling off in larger groups.

MAC, whose efficiency is lowest among the six top strategies at group sizes of two and three, experiences a much sharper rate of increase than MAE. Moreover, MAC continues to increase more sharply than MAE, SHU and FRI, even after assuming the lead at the group size of seven. The larger the competing population, the better MAC performs, relative both to its own increasing efficiency and to the efficiencies of its competitors.

ELIMINATORY ECOSYSTEMIC COMPETITION

Next, the strategic population is subjected to a different measure of robustness; namely, an ecological scenario.

The ecological scenario emerges as an offshoot of evolutionary game theory.[5] Following Axelrod's and Hamilton's (1981) attempt to find an evolutionarily stable strategy for iterated prisoner's dilemmas, Boyd and Lorberbaum (1987), Axelrod and Dion (1988) and Marinoff (1990) have all shown, in different ways, that the prisoner's dilemma is not susceptible to evolutionary modelling in the original Maynard Smith sense. The ecological scenario, however, provides an interesting alternative perspective on strategic robustness. Axelrod's (1980b) scenario simulates "survival of the fittest," by assuming that the strategies' scores represent relative numbers of offspring, which continue to compete in subsequent "generations" of the tournament.

5. E.g. see Maynard Smith (1982).

TABLE 6
Overall Efficiencies: 524,287 Appearances in 616,666 Subtournaments

	Number of Competitors									
	2	3	4	5	6	7	8	9	10	11
MAC	52.6	61.7	68.7	74.0	78.0	81.3	83.9	86.2	88.2	90.1
MAE	78.9	78.9	79.5	79.3	79.5	79.8	80.3	80.8	81.3	81.8
SHU	100	84.5	78.2	76.3	76.3	77.2	78.4	79.6	80.8	81.8
FRI	100	87.4	81.3	78.5	77.5	77.2	77.3	77.4	77.5	77.7
CHA	73.7	67.3	67.3	69.1	71.2	73.0	74.4	75.6	76.6	77.4
ETH	73.7	71.6	69.6	70.0	71.0	72.1	73.1	73.9	74.7	75.3
TFT	100	80.4	74.3	70.7	68.9	68.2	67.9	67.9	67.9	68.0
TES	68.4	74.6	73.3	71.2	70.0	69.4	68.9	68.5	68.1	67.7
MEU	63.2	68.7	68.1	67.5	67.1	66.7	66.5	66.4	66.4	66.5
TTT	78.9	64.9	59.4	57.7	57.0	57.0	57.1	57.1	57.1	57.1
MAD	42.1	42.4	41.1	41.8	40.5	40.4	40.1	39.3	39.3	38.7
GRO	63.2	49.1	42.8	39.9	40.8	40.2	39.2	39.0	38.5	38.2
TQD	36.8	41.8	40.6	39.9	39.4	39.0	38.5	38.0	37.8	37.4
BBE	5.3	13.5	20.4	24.9	27.9	29.6	30.9	31.9	32.5	33.3
DDD	42.1	36.8	36.8	35.1	32.8	31.8	30.5	29.3	28.3	27.5
RAN	42.1	37.1	35.2	33.6	32.1	30.6	29.5	28.5	27.7	27.0
TAT	42.1	31.6	30.6	29.0	27.2	25.6	24.3	23.0	21.8	20.8
NYD	52.6	30.4	21.6	18.6	17.2	16.5	16.2	15.7	15.3	15.0
TQC	31.6	28.4	21.8	19.3	16.7	14.4	13.5	12.8	11.7	11.1
CCC	47.4	25.7	18.0	13.4	11.9	10.9	9.8	9.2	8.5	7.9

A straightforward mathematical notation is introduced in order to show explicitly how the interactive ecological algorithm functions.

$(AvB)_n$ means "the relative number of n^{th} generation offspring produced by A-strategists in competition against B-strategists." Thus $(AvB)_1$ is strategy A's tournament score against strategy B.

$(TOT)_{A,\,n}$ means "the total relative number of n^{th} generation offspring produced by A-strategists in competition against all strategists"; that is:

$$(TOT)_{A,\,n} = (AvA)_n + (AvB)_n + \ldots + (AvZ)_n$$

for Z different strategies in the environment. Thus $(TOT)_{A,\,n}$ is strategy A's total score in the tournament.

The relative frequency of A-strategists in the initial generation is the ratio of strategy A's total offensive score ($[TOT]_{A,\,1}$) to the sum of all strategies' total offensive scores. In the initial generation, all relative frequencies are

Number of Competitors									
12	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	E_0
91.8	93.6	95.2	96.6	97.9	98.9	99.6	100	100	88.8
82.4	83.0	83.6	84.3	85.1	86.1	87.8	90.1	90.0	81.6
82.8	83.6	84.5	85.4	86.4	87.4	88.6	89.8	85.0	81.2
77.8	78.0	78.3	78.6	78.9	79.4	80.2	81.6	80.0	77.7
78.1	78.8	79.4	80.0	80.4	80.5	80.3	79.5	75.0	76.7
75.7	76.1	76.3	76.4	76.3	76.0	75.4	74.3	70.0	74.7
67.9	67.8	67.5	67.2	66.8	66.5	66.3	65.8	60.0	67.9
67.3	66.8	66.2	65.7	64.9	64.0	62.6	60.8	55.0	67.8
66.5	66.6	66.7	66.7	66.7	66.6	66.6	66.7	65.0	66.5
57.0	56.9	56.8	56.6	56.5	56.2	55.7	54.7	50.0	57.0
38.2	38.2	37.4	37.3	37.1	36.3	37.7	36.5	30.0	38.9
38.0	37.5	37.2	37.4	37.4	38.5	39.7	40.1	40.0	38.4
37.2	37.0	36.8	36.8	36.9	36.8	36.8	37.4	35.0	37.7
34.2	35.1	36.1	37.1	28.0	38.7	39.8	43.0	45.0	32.9
26.5	25.7	25.3	24.0	23.5	23.7	22.4	22.2	25.0	28.1
26.2	25.7	25.4	24.9	25.0	25.3	24.3	24.6	20.0	27.6
19.8	18.8	18.1	17.5	16.7	15.7	15.8	15.2	15.0	21.5
14.6	14.4	14.1	13.8	13.4	12.9	12.8	11.4	10.0	15.2
10.6	9.6	8.7	8.0	6.8	6.1	4.7	5.6	5.0	11.6
7.4	6.9	6.5	5.8	5.3	4.1	3.2	0.9	0.0	8.4

computed directly from the tournament matrix of raw scores. Then, for each subsequent generation, the recurrence relation

$$(A\text{v}B)_{n+1} = (A\text{v}B)_n[(TOT)_{A,n}]/[(TOT)_{A,n} + (TOT)_{B,n}]$$

is used to compute the new matrix of offspring, from which each strategy's relative frequency in that generation can be found. This algorithm is applied to all strategies in the environment, and is iterated for a sufficient number of generations, until all rates of growth (and decline) subside to a quiescent state.

Note that if the n^{th} generation ratio of offspring, $([A\text{v}B]_n)/([B\text{v}A]_n)$, has the numerical value a/b, then the $(n + 1)^{st}$ generation ratio will be

$$[(A\text{v}B)_{n+1}]/[(B\text{v}A)_{n+1}] = (a/b)[(TOT)_{A,n}]/[(TOT)_{B,n}]$$

This satisfies the two principal requirements of Axelrod's model; namely, that the ratio of offspring between two competing strategies in any future

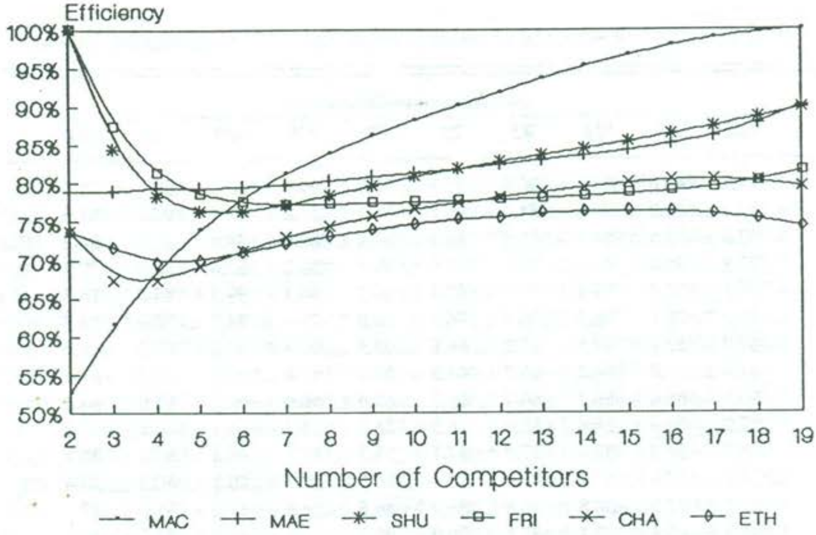

Figure 1: Most Efficient Combinatoric Performances: All Subtournaments, 2-19 Competitors

generation be proportional to: (a) their ratio of offspring in the previous generation, and (b) their relative frequencies in the previous generation.

The relative frequency of each strategy's progeny in a given generation is expressed in parts per thousand (*ppt*) of the overall population in that generation.

The ecological scenario involving the 20 strategies of the interactive tournament attains a stable state after about 325 generations. That is to say, following the 325th generation, the rate of change has slowed to the extent that all strategies' cumulative increases or decreases in relative frequency are less than one *ppt* over the next several generations. Although minor fluctuations continue to take place in increments (or decrements) of parts per ten thousand per generation and less, these fluctuations are negligible on the scale of the scenario.

The results of the ecological scenario involving 20 strategies are displayed in Figure 2, which shows the initial (parent generation) and stable (325th generation) frequencies for each strategy. The strategies appear, from left to right, in descending order of their stable frequencies.

It is clear from Figure 2 that MAC, which has the largest initial frequency (60 *ppt*), experiences the greatest increase, to a stable frequency of 142 *ppt*.

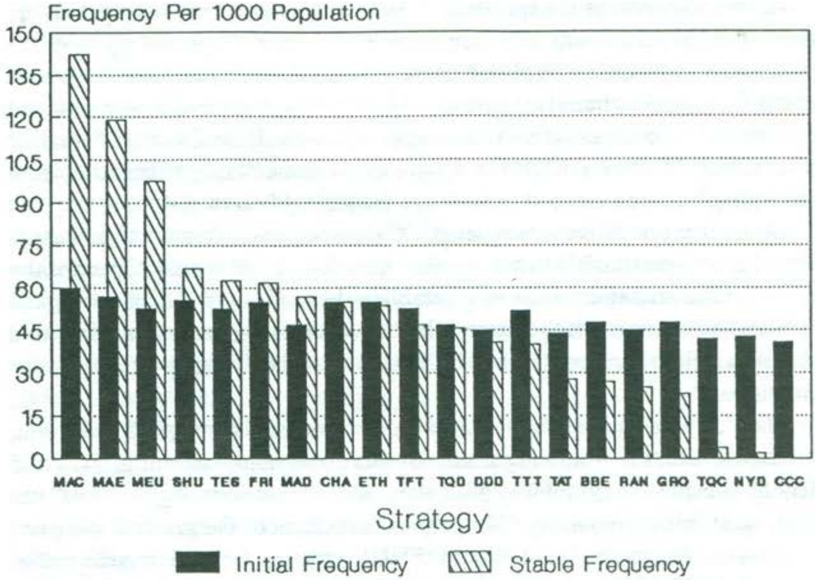

Figure 2: Ecology of Main Tournament: Initial versus Stable States

This represents an increase of 82 *ppt* over 325 generations, or an average growth rate of 0.25 *ppt* per generation. And MAE, which has the second largest initial frequency (57 *ppt*), experiences the second greatest increase, to a stable frequency of 119 *ppt*. MAE's average rate of growth is thus 0.19 *ppt* per generation.

Obviously, the size-order of the initial frequencies is identical to the rank order of the tournament, because a strategy's initial frequency is its total tournament score divided by the sum of all strategies' total tournament scores, and this dividend remains constant (for a given matrix of raw scores). However, the size-order of the stable frequencies does not necessarily correspond to that of the initial frequencies. For example, SHU ranks third in initial frequency (56 *ppt*), but slips to a distant fourth in stable frequency (67 *ppt*). SHU is overtaken by MEU, which ranks only seventh in initial frequency (53 *ppt*), but third in stable frequency (98 *ppt*). SHU's growth rate is 0.034 *ppt* per generation; MEU's is 0.14 *ppt* per generation.

That MAC, MAE and MEU produce the greatest relative numbers of progeny, respectively, is a testament not only to their individual fitnesses, but also to the overall fitness of the maximization family in this ecosystem.

At the other end of the spectrum, it is probably not a coincidence that the greatest declines in frequency are experienced by CCC (– 41 *ppt*), NYD (– 41 *ppt*), and TQC (– 38 *ppt*). Not only do these strategies have the lowest initial frequencies, but also, perhaps significantly, their order of ecological decline corresponds exactly to their order of points allowed in the interactive tournament. Moreover, CCC has apparently become "extinct" because, from the 10th generation onward, its relative frequency is zero *ppt*.

Given that one of the strategies (CCC) has become extinct in this ecosystem, it seems reasonable to ask another question: What would happen if the 19 surviving strategies were to reestablish themselves in a new ecological habitat, beginning with the same initial conditions and subject to the same generative algorithm save that all CCC-strategists have disappeared from the population?

The scenario is thus regenerated in a new ecosystem of 19 surviving strategies, with the following result. In this ecosystem, rates of growth and decline subside to negligibility after about 450 generations. Again, MAC has the largest initial frequency (62 *ppt*), and experiences the greatest increase, to a stable frequency of 176 *ppt*. SHU, with the second largest initial frequency (60 *ppt*), ranks fourth at stability (98 *ppt*). MAE, which has the sixth largest initial frequency (58 *ppt*), vaults past FRI, CHA, ETH, and SHU to rank second at stability (138 *ppt*). And MEU, initially in a three-way tie for eighth place (55 *ppt*), finishes third at stability (99 *ppt*). The maximization family continues to exhibit reproductive fitness in this ecosystem.

This procreative model is clearly sensitive to perturbation (by the removal or, inversely, by the addition of a competing strategy). The term "ecology" seems well-chosen by Axelrod in that the extinction of one strategy has palpable repercussions on the interactions among the 19 survivors. In the original ecosystem, both MAE and MAC enjoy comparatively high reproductive success in competition with CCC. As soon as CCC becomes extinct, MAE falls from second to fourth place in initial frequency; MEU, from sole possession of seventh to a three-way tie for eighth. That MAE and MEU now overtake numerous competitors in order to finish second and third overall behind MAC, illustrates their fitness in regaining lost reproductive ground.

The perturbation also results in the extinction of two more strategies: in this new ecosystem, NYD's progeny vanish after the 10th generation; TQC's, after the 11th. Once again, the first strategy to become extinct in this ecosystem is the strategy with the lowest initial frequency (NYD, 41 *ppt*). But TQC, which disappears one generation later, shares the second-lowest initial frequency with TAT (43 *ppt*). Although TAT experiences a sharp decline, it manages to stabilize at 8 *ppt*.

The eliminatory process is continued by establishing another ecosystem, composed of the eighteen surviving strategies after the demise of NYD. This ecosystem is similarly procreated until stability is attained, whereupon another new ecosystem is formed by deleting the next strategy to become extinct. This eliminatory process is repeated to its eventual conclusion.

Not surprisingly, the number of generations required to stabilize an ecosystem is not a smoothly decreasing function of the number of competing strategies. Although such a trend is observable overall, many individual reversals of that trend occur. Because the number of generations required to attain stability diminishes only in tendency with the number of competing strategies and seems to depend critically on the particular combination of strategies in competition, one can conclude that this eliminatory process is somewhat stochastic.[6]

In the ecosystem involving 18 strategies (following the extinction of NYD), SHU holds the greatest initial frequency (65 *ppt*), and MAC and FRI are tied with the second greatest (64 *ppt*). FRI experiences the largest increase, however, and realizes the greatest stable frequency (154 *ppt*), followed by SHU (152 *ppt*) and MAC (144 *ppt*). MAE initially ranks seventh (60 *ppt*), but climbs to fourth at stability (114 *ppt*), while MEU initially ranks ninth (56 *ppt*) but finishes fifth (79 *ppt*). Thus MAC, MAE, and MEU continue to perform quite well, but they slip to third, fourth, and fifth places with respect to magnitudes of stable frequencies.

A glance at the tournament matrix of raw scores (Table 2) affords an explanation for what is taking place. In the context of the tournament, the maximization family fared extremely well against CCC and NYD. In fact, each member of the maximization family realizes its two highest scores against these very strategies. But in the ecological context, this large margin of success not only contributes to the rapid extinction of the weaker strategies, but also proves detrimental to the exploitive ones.

In the tournament, for example, MAC out-scored CCC by 4,824 to 264. So in the ecological scenario, their parent generation ratio is thus 4,824:264, or about 18:1 in favour of MAC. And in the parent generation of the 20 strategy ecosystem, their respective initial frequencies are 60 and 41 *ppt* of the overall population. Thus the ratio of their second-generation offspring is

6. An exponential curve fit, which gives the number of generations required to stabilize the population frequencies as a function of the number of strategies in competition yields the result $y = 26e^{.138x}$ with an unimpressive correlation $r = .82$. However, this equation offers one possible explanation as to why Axelrod's ecosystem does not attain stability after 1,000 generations. With 63 competitors, the equation predicts that 155,000 generations are required to attain stability. Of course, any such extrapolation remains highly conjectural.

(4,824 × 60):(264 × 41), or about 27:1 in favour of MAC. In the tournament context, MAC exploits CCC rather heavily (as do many other strategies) with no dire consequences to itself. But in the ecological context, MAC's heavy exploitation of CCC has a three-fold result.

First, MAC benefits from a proportionately large increase in progeny. Second, CCC, which experiences a generally poor differential procreative rate in the ecosystem as a whole, is unable to stave off elimination. Third, in subsequent ecosystems, MAC no longer benefits from its high procreative rate in competition against CCC, because CCC is extinct. In future ecosystems, MAC must compete more frequently against strategies with greater procreative fitness than CCC, strategies that MAC cannot exploit as readily.

This is a classic instance of overexploitation of a resource to the eventual detriment of the exploiters. All strategies that overexploit CCC (such as DDD, TQD, TAT, TES, and the maximization family) abet CCC's rapid extinction and, in so doing, deprive themselves of a competitor that allows them to create large relative numbers of progeny. When a new ecosystem is established with CCC absent from the environment, the population frequencies undergo an ecological shift such that those strategies that overexploited the extinct competitor now experience corresponding declines in their procreative rates. In future ecosystems, former exploiters may themselves become the victims of exploitation.

In the ecosystem with 17 strategies (following the extinction of TQC), the stable order is once again FRI (136 *ppt*), SHU (130 *ppt*), MAC (129 *ppt*), MAE (108 *ppt*), and MEU (86 *ppt*). The population gaps between these upper five strategies have closed, compared with the previous ecosystem. And now, with TAT's extinction, one observes that, in the first four ecosystems, the lower four strategies of the tournament have become extinct in reverse order of their tournament ranks from 20th to 17th (*CCC, TQC, NYD, TAT*).

TAT's extinction (combined with the previous extinctions) results in a reordering of initial frequencies in the next ecosystem, which precipitates new stable standings. In the ecosystem with 16 strategies, CHA (76 *ppt*), ETH (75 *ppt*), and TFT (71 *ppt*) are most successful, both initially and at stability, realizing eventual frequencies of 125, 122, and 102 *ppt* respectively. TTT places fourth at stability (95 *ppt*), and SHU manages a tie for fifth with TES (93 *ppt*). Evidently, TAT's extinction results in a complete upheaval in the environment with new strategies in the ascendancy and previously successful strategies in decline. MAC slips to seventh at stability (86 *ppt*); MAE, ninth (66 *ppt*). Moreover, this ecosystemic competition requires the greatest number of generations (more than five hundred) to settle down. In addition, the precedent for extinction is broken. DDD (which ranks ahead of RAN in the tournament) now vanishes from the ecology.

In ecosystems involving from 15 to 10 competitors, CHA and ETH continue to predominate at stability, while MAC, SHU, TFT, and TTT also tend to flourish. In ecosystems involving from 9 to 5 competitors, TFT ranks first four times and second once. The ecosystemic competition of seven strategies is won by TES. In this competition, TES experiences the greatest increase of any strategy in any ecosystem—from an initial frequency of 150 *ppt* to a stable frequency of 397 *ppt* after 36 generations. But TES becomes extinct in the ecosystem of 5 competitors.

The final ecosystem is composed of four nice strategies: TFT, SHU, GRO, and ETH. In such a system, all future generations of progeny maintain respective ratios of 1:1. Thus, initial frequencies and stable frequencies are identical and equal to one another, and stability is attained in the parent generation. This situation would, of course, obtain in an ecosystem of any size, providing that it were composed exclusively of nice strategies. The other nice strategies, however (namely CHA, TTT, FRI, NYD and CCC), are already extinct, because their respective combinations of attributes were disfavoured in previous ecosystemic competitions.

The nine earliest extinct strategies have aggregate decreases in frequency, but three of the last four to become extinct, as well as one of the survivors, also have aggregate decreases. Thus, although an aggregate increase in frequency indicates that a competitor does not face early extinction, neither is it a passport to ultimate survival.

One might find the survival of GRO perplexing. GRO experiences an increase in only 2 of the 16 ecosystemic competitions that result in an extinction; nonetheless GRO survives to the final ecosystem. GRO does not excel in any of these competitions and ranks near the bottom in all of them. Yet GRO is tenacious enough to survive them all, apparently by dint of consistent mediocrity. Because GRO is never highly successful, it cannot be said to depend on any particular strategies for its success. Hence GRO is not subject to the vicissitudes of overexploitation, which cause the rise and fall of many of its more successful, and later extinct, competitors.

Similarly, TES is the last strategy to become extinct. Despite its superlative performance in the one ecosystemic competition, TES also has an aggregate decrease in frequency.

Then again, one finds that ETH has the largest aggregate increase in frequency, but ETH wins only two of the eliminatory ecosystemic competitions. Moreover, CHA has a larger aggregate increase than three of the four survivors, yet CHA eventually succumbs to extinction.

TABLE 7
Top Five Strategies, With Respect to Stable Frequency

Competitors	First Place	Second Place	Third Place	Fourth Place	Fifth Place
20	MAC	MAE	MEU	SHU	TES
19	MAC	MAE	MEU	SHU	FRI
18	FRI	SHU	MAC	MAE	MEU
17	FRI	SHU	MAC	MAE	MEU
16	CHA	ETH	TFT	TTT	SHU
15	CHA	ETH	TFT	SHU	TTT
14	CHA	ETH	MAC,SHU	–	TFT
13	CHA	MAC	SHU	ETH	TFT
12	CHA	ETH	MAC	TTT	TFT
11	ETH	CHA	TTT	TFT	TES
10	ETH	CHA	TTT	GRO	MAC
9	TFT	ETH	TES	TTT	CHA
8	TFT,ETH,SHU	–	–	GRO	CHA
7	TES	TFT,SHU,ETH	–	–	GRO
6	TFT,SHU,ETH	–	–	TES	GRO
5	TFT,SHU,ETH	–	–	GRO	TES

Table 7 illustrates the waxing and waning fortunes of the top five ranking strategies at stability for each of the ecosystemic competitions.

Table 7 shows that, in general, the maximization family is most successful in the larger ecosystems; the optimization family, in the medium-sized ecosystems; and the tit-for-tat family, in the smaller ecosystems. But no single strategy emerges as most robust overall if the sole criterion of robustness is stable frequency. Indeed, although several strategies claim varying degrees of success in different sizes of ecosystem, it does not seem possible to ascribe a coherent order of robustness from one criterion alone.

One criterion suffices for Axelrod, who conducts a single ecosystemic competition among 63 strategies. Based solely on its magnitude of relative frequency in the population, TFT wins that particular competition. However, given what transpires in eliminatory ecosystemic competitions in the environment of the interactive tournament, it seems reasonable to speculate that, if a similar range of competitions were conducted in Axelrod's environment, no single strategy would win them all. It seems rather more likely that one would observe a similar waxing and waning of strategic procreativity in different ecosystems.

ECOLOGICAL ROBUSTNESS

The question remains: How can one assess robustness across the range of ecosystemic competitions? Clearly, there is no unique way to accomplish this task. One possible method consists in a parametric approach. The parameters themselves are quantifications of vital attributes of robustness in the ecological context. In other words, the above question is answered in three stages. First, vital properties of an ideal ecologically robust strategy are posited. Second, the varying extents to which the competing strategies embody these properties are quantified according to appropriate ranking schemes. Third, these quantifications are enlisted as parameters that reflect each strategy's combined embodiment of vital properties and that permit a corresponding overall index of robustness to be assigned.

This experiment uses four parameters drawn from the ecological scenario. Four vital properties of an ecologically robust strategy are posited and their corresponding parameters are defined as follows:

1. The ideal ecologically robust strategy's progeny are able to avoid extinction. Hence the first parameter is survival or ecosystemic longevity. Each strategy is ranked in ascending order of the total number of generations during which its progeny avoids extinction.
2. The ideal ecologically robust strategy is reproductively fit; that is, its number of progeny increases in future generations. Hence the second parameter is overall average increase in relative population frequency between initial and stable states of every ecosystem. Each strategy is ranked in ascending order of the quotient of its aggregate frequency and the number of generations its progeny survives. This quotient is thus a measure of a strategy's average increase in relative frequency, in parts per thousand of the population per generation extant. (A negative increase, of course, indicates a decrease.)
3. The ideal ecologically robust strategy maintains a consistently high stable frequency from one ecosystemic competition to another. Hence the third parameter is overall stable efficiency. A strategy's stable efficiency is computed in the following way. Suppose that strategy A has the j^{th} -largest stable frequency in an ecosystemic competition involving k competitors (including itself). Thus, strategy A achieves a higher stable frequency than $(k - j)$ other competitors. Its best possible performance (if it finishes first) entails achieving a higher stable frequency than $(k - 1)$ other competitors (excluding itself). Hence, strategy A's relative stable efficiency in this competition is $(k - j)/(k - 1)$. Strategy A's overall relative stable efficiency, in n ecosystemic competitions, is therefore

$$\sum_i (k_i - j_i)/(k_i - 1),$$

which is the net ratio of the number of competitors it betters to the number of competitors it faces. Each strategy is ranked in ascending order of its overall stable efficiency.

4. The ideal ecologically robust strategy shows adaptivity across the range of ecosystemic competitions, by means of consistent improvement within them. That is, it consistently increases its frequency, relative to other competitors, thereby tending to improve its position in a given competition. Hence the fourth parameter is the sum of the fractions of competitors overtaken in each competition divided by the total number of competitions. If a strategy overtakes j_1/k_1 competitors in its first competition, j_2/k_2 competitors in its second competition, and so on, up to and including j_n/k_n competitors in its n^{th} competition, then that strategy's average adaptivity is:

$$(1/n)\sum_i j_i/k_i.$$

Each strategy is ranked in ascending order of the signed magnitude of its adaptivity whose dimensions are: average fraction of competitors overtaken per competition. (A negative adaptivity obtains when a strategy is overtaken by more competitors than it overtakes.)

Now one has four different ranking schemes that order the strategies in terms of the four parameters: longevity, fecundity, stability, and adaptivity. These rank parameters are abbreviated respectively as R_l, R_f, R_s, and R_a. With each strategy is then associated a unique set of four rank numbers that correspond to that strategy's particular values for $\{R_l, R_f, R_s, R_a\}$.

A given strategy's index of robustness, I_r, is evaluated in the following way. Each of its four rank numbers is subtracted from 20 to give the number of competitors it betters according to each parameter. These four new numbers are then added and their sum is divided by 76 (which is the total number of competitors it could have bettered overall; i.e., 19 competitors in each of four schemes). This normalized quotient is the given strategy's index of robustness. That is,

$$I_r = [80 - (R_l + R_f + R_s + R_a)]/76.$$

The ideal ecologically robust strategy would rank first in each scheme, and its index of robustness would then attain the maximum value of unity. An utterly nonrobust strategy would rank 20th in each scheme, and its index of robustness would take on the minimum value of zero.

The magnitudes of the four parameters, their corresponding rank numbers, and the resulting indices of robustness are displayed in Table 8.

TABLE 8

Four-Parameter Quantification of Ecological Robustness

	Longevity	R_l	*Average Fecundity*	R_f	*Stable Eff%*	R_s	*Adaptivity*	R_a	I_r
MAC	2995	9	+.117	3	77.2	2	+.135	3	.829
SHU	3261	1	+.109	4	75.9	3	–.025	11	.803
ETH	3261	1	+.232	1	80.2	1	–.089	17	.789
TFT	3261	1	+.108	5	72.7	5	–.062	14	.724
MAE	2666	11	+.059	6	61.4	6	+.128	4	.697
MEU	2447	12	–.015	7	51.9	10	+.137	2	.645
CHA	3227	6	+.150	2	74.4	4	–.236	19	.645
FRI	3106	8	–.042	10	58.0	7	+.038	7	.632
MAD	1789	15	–.063	11	37.4	11	+.181	1	.553
TES	3260	5	–.023	8	56.5	8	–.157	18	.539
TTT	3179	7	–.038	9	53.7	9	–.074	15	.526
DDD	1259	16	–.085	12	27.1	13	+.116	5	.447
GRO	3261	1	–.159	16	27.3	12	–.271	20	.408
TQD	2031	14	–.102	13	25.0	15.	+.025	8	.395
BBE	2917	10	–.135	14	26.0	14	–.034	12	.395
TAT	795	17	–.172	17	12.9	16	+.068	6	.316
RAN	2383	13	–.145	15	11.3	17	–.061	13	.289
TQC	345	18	–.365	19	7.4	18	+.018	9	.211
CCC	9	20	–4.60	20	0.0	20	±.000	10	.132
NYD	335	19	–.257	18	5.4	19	–0.82	16	.105

According to this parametric approach, MAC is the most ecologically robust strategy, followed by SHU, ETH, TFT and MAE, to round out the top five. Although MAC became extinct earlier than its most robust rivals (which rank first in longevity compared to MAC's ninth), these rivals prove comparatively less adaptive. In fact, SHU, ETH and TFT are all negatively adapted; that is, they are surpassed, on average, by a larger fraction of competitors than they surpass.

These parameters are quite revealing with respect to the competitive performance of a given strategy, as indeed they must be if they are to provide a reasonable quantification of ecological robustness.

Examine the case of ETH, for example. ETH shares the greatest longevity, produces the largest average number of offspring per generation, and is most efficient in overall stable frequency rankings. Given this outstanding combination of attributes, one might expect ETH to win a substantial number of ecosystemic competitions. Yet a glance at Table 8 shows that ETH is the outright winner in only two of the competitions. Moreover, those competitions do not involve a relatively large number of strategies (11 and 10 strategies, respectively). In fact, in Table 8, ETH is conspicuously absent from the top rankings in competitions involving 20, 19, 18 and 17 strategies. Why does ETH not fare better?

The fourth parameter provides an explanation. ETH turns out to be one of the least adaptive strategies. ETH's great longevity, prodigious fecundity, and high efficiency do not reveal its principal weakness: In larger groups, ETH is readily overtaken by a substantial fraction of competitors. These competitors, which produce fewer progeny on average than ETH, and which better fewer strategies overall than ETH, are nevertheless more reproductively fit than ETH when the competitive traffic is heaviest. Thus, notwithstanding ETH's fortitude with respect to three attributes, ETH's robustness is compromised by an acute lack of adaptivity in large groups.

The seven most robust strategies, not surprisingly, are also the seven most fecund (although not in that order). The three most robust strategies are also the most efficient (although again, not in that order). Overall, fecundity and efficiency are the most closely correlated pair of attributes. But the two most robust strategies, MAC and SHU, show respective improvements in rank with respect to this attribute pair. MAC ranks third in fecundity and second in efficiency; SHU, forth in fecundity and third in efficiency. This type of improvement, however slight, denotes an interesting performance characteristic—namely, an effective frequency distribution of progeny across the range of ecosystemic competitions.

Average relative frequencies, by definition, do not take instantaneous changes in frequency (from one competition to another) into account. MAC experiences an increase in progeny in 9 of its 12 competitions; SHU experiences an increase in 11 of its 17 competitions. Both MAC and SHU achieve frequency distributions which, in terms of rank efficiency, enable these strategies to realize the beneficial potential of their increases and to minimize the detrimental effects of their decreases.

In contrast, CHA ranks second in fecundity, but slips to fourth in efficiency. Although CHA's average increase in progeny is greater than that of MAC and SHU, CHA's distribution of instantaneous increases is less effective. CHA experiences an increase in progeny in 10 of its 15 competitions (and no change in one competition), but its largest increases occur in competitions in which a smaller increase would confer the same efficiency rank. In other words, CHA produces more offspring than it requires in some situations and not enough in others. CHA is nonetheless relatively robust, although its robustness is severely compromised by its poor adaptivity.

The point to be made here is that, notwithstanding instances of pair-wise correspondence between R_f and R_s, these two rank parameters reflect quite distinct attributes. A given strategy's difference in rank between these parameters (or lack thereof) is indicative of a particular performance characteristic.

Table 9 compares strategic robustness in the combinatorial subtournaments with strategic robustness in this ecological scenario. The order of overall robustness is determined by taking the average of each strategy's rank with respect to combinatorial and ecological robustness. Because MAC ranks first in both categories, it is obviously most robust overall in the interactive environment. SHU is deserving of second overall, and MAE retains third overall despite its decline in the ecological scenario.

Once again, it must be stressed that this parametric approach to the evaluation of ecological robustness is by no means a unique determinant; many other schemes could be conceived and applied. The addition or deletion of a single parameter can alter the standings, either mildly or drastically. Although more (or fewer) than four parameters could be used, the result in this case seems reasonably unbiased. At the least, an attempt has been made to neutralize or otherwise balance any bias that inheres in such a quantification.

The ecological scenario is clearly rich in interactions and implications, and many more such models can and should be developed within the evolutionary paradigm. The main difference between ecological and evolutionary modelling is, as Axelrod (1980b) points out, that the former does not admit of any "mutational" influences. In other words, the ecology unfolds

TABLE 9
Overall Strategic Robustness

Rank	Combinatorial Robustness	Ecological Robustness	Overall Robustness
1	MAC	MAC	MAC
2	MAE	SHU	SHU
3	SHU	ETH	MAE
4	FRI	TFT	ETH
5	CHA	MAE	TFT, CHA
6	ETH	MEU, CHA	–
7	TFT	–	FRI
8	TES	FRI	MEU
9	MEU	MAD	TES
10	TTT	TES	MAD
11	MAD	TTT	TTT
12	GRO	DDD	GRO
13	TQD	GRO	DDD, TQD
14	BBE	TQD	–
15	DDD	BBE	BBE
16	RAN	TAT	RAN, TAT
17	TAT	RAN	–
18	NYD	TQC	TQC
19	TQC	CCC	NYD
20	CCC	NYD	CCC

strictly from initial conditions with no behavioral modifications made to the strategies involved. However, it is evident that, on the basis of strategic interaction alone, and in the absence of strategic modification, the complexities of eliminatory ecosystemic competition necessitate correspondingly complex methods of assessing robustness.

CONCLUSIONS

The results of the interactive tournament provide a number of empirical corroborations of, as well as certain departures from, Axelrod's principal findings. Five conclusions are briefly presented here.

First, one observes the relative success of cooperative maximization rules. Following his first tournament, Axelrod showed that, had DOWNING's initial view of its opponent's responsiveness been more optimistic, then DOWNING would have won and won by a large margin. DOWNING is a

TABLE 10
The Costs of Exploitation

	Versus CCC	*Versus TFT*	*Average 1/2 (CCC + TFT)*	*Tournament Average*
MAC	4826	2970	3898	2645
MAE	4848	1272	3060	2503
MEU	4900	1148	3024	2362
MAD	4986	1041	3014	2086

reactive maximization strategy whose initial probabilistic weights are identical to those of MEU. The interactive tournament provides a clear indication that the success of a maximization rule against a broad spectrum strategic population increases strictly with its propensity to cooperate. In terms of combinatorial, ecological and overall robustness, MAD, MEU, MAE, and MAC finished in ascending order (10th, 8th, 3rd and 1st) of their initial cooperative weightings (1/10, 1/2, 5/7, and 9/10), respectively.

Second, one can specify a single condition that accounts for the relatively strong performances of MAC and MAE (and the relatively weak showings by MEU and MAD). Following his second tournament (which contained a field of 62 entries), Axelrod concluded that "being able to exploit the exploitable without paying too high a cost with the others is a task which was not successfully accomplished by any of the entries." In the interactive tournament, MAC is able to accomplish this task to great advantage.

Table 10 illustrates the fortunes of the maximization family against an utterly exploitable strategy (CCC), and against a highly provocable (and therefore nonexploitable) strategy (TFT). MAD, MEU, and MAE are able to exploit CCC to slightly greater extents than MAC. But MAD, MEU, and MAE all incur a vastly higher cost against TFT, whereas the MAC-TFT pair attains a score (tied at 2,970 points) comparable to that attained by two nice strategies (tied at 3,000 points). MAC's advantage lies in its overwhelmingly cooperative weighting, which allows it to accomplish Axelrod's formerly hypothetical task.

Third, Axelrod found niceness to be a key ingredient of successful performance. In the interactive environment, nide strategies placed first and third in terms of overall robustness. (Recall, a nide strategy is neither nice nor rude; i.e., is indeterminate with respect to primacy of defection.) So, although rudeness again leads to failure, one finds that nideness – as opposed to niceness – along with provocability, forgiveness, and exploitiveness, can

TABLE 11
The Maximization Family versus Itself

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

also conduce to success.

The performances of the most robust members of the tit-for-tat family, namely SHU, TFT and TTT, afford an insight into the cause of this shift in criteria. Axelrod (1980a) asserts that, had TTT been entered in his first tournament, it would have won. In that case, the order of finish of these members would have been: TTT, TFT, SHU. Note that the ascending order is one of strictly increasing forgiveness. However, in the interactive tournament, these strategies' ascending order of finish is exactly reversed: SHU, TFT, TTT. This order is one of strictly decreasing forgiveness. From this comparison, one can infer that the environment of the interactive tournament is much harsher than that of Axelrod's first tournament. In a friendly environment, one would naturally expect forgiveness to conduce to success; in a harsh environment, one would just as naturally expect a lack thereof to succeed. Thus one can conclude that nideness is not necessarily preferential to niceness; nideness merely supersedes niceness under certain conditions.

And a caveat to the maximization family's particular brand of nideness is expressed in a rather significant phenomenon: A maximization family member always reacts to an occurrence of joint cooperation by becoming nice instead of nide. If one instance of mutual cooperation occurs, a maximization strategy will never be the first to defect thereafter.[7]

Fourth, Axelrod argued cogently that, notwithstanding its victories in both his tournaments, TFT is not the "best" rule for iterated prisoner's dilemmas. Axelrod gave examples of rules that would have won his first tournament, had they been submitted. And the interactive tournament affords an example of a competitive population in which TFT is surpassed by a number of other strategies; some hitherto untried, others present in previous tournaments.

7. In terms of this family's decision calculus, every instance of mutual defection has the effect of *lowering* the value of the expected utility of further defection. Thus mutual defection actually increases the maximization strategy's propensity to cooperate.

Axelrod's overriding observation is that "there is no best rule indepen-dent of environment." The case of MAC brooks no exception, and contains an irony well worth noting. If one seeks an environment in which MAC is not the best rule – indeed, in which MAC is next-to-worst – then one need look no further than the environment of the maximization family itself. As Table 11 illustrates, MAC performs dismally against its own siblings and twin.

Fifth, the most general reason for MAC's undistinguished intrafamilial performance is indirectly cited by Axelrod (1980b, 401-2), albeit in a different context: "TIT FOR TAT could have been beaten in the second round [tournament] by any rule which was able to identify and never cooperate with RANDOM, while not mistaking other rules for RANDOM." Indeed, an examination of the maximization family's scores against the probabilistic family (see Table 3) reveals that the maximization calculus readily identifies strategies that play randomly (regardless of their weightings), and prescribes perpetual defection against them during the deterministic phase of its play. At the same time, however, maximization family members exhibit precious little recognition of one another. Owing to their random play during the first one hundred moves, most members naturally but mistakenly identify their familial opponents as random strategies, and proceed to defect perpetually against them as well.

Only the MEU-MEU and MAE-MAE twins seem able to achieve eventual perpetual mutual cooperation. This begs a vital question: Why do the MAC-MAC twins, whose initial mutual cooperativeness is probabilistically the highest in the family, fail to achieve perpetual mutual cooperation even sooner than their less cooperatively-weighted siblings?

The complete answer to this ironic question is not trivial. It lies in a detailed investigation of the intricate and counterintuitive properties of the event matrix itself. Such an exposition is reserved for a future report. Suffice to say that the more cooperative its weighting, the less reliable is the performance of a maximization strategy against its twin.

Despite its intrafamilial difficulties, the cooperatively weighted maximization of expected utility is demonstrably robust in the relatively harsh environment of the interactive tournament. Accumulated evidence suggests that it could have won both of Axelrod's tournaments as well.[8] MAC is nide, forgiving, provocable, and exploitive. It also has the capacity to become nice. Finally, MAC is unrelievedly optimistic, for it is able to enlist mutual defection in the service of perpetual mutual cooperation (see note 7). It

8. In shorter games, such as those conducted by Axelrod, MAC's initial number of random moves would be reduced proportionately.

therefore embodies the precept that, in certain circumstances, the game-theoretic end may well justify the game-theoretic means.

A final word of caution: Owing to MAC's exploitiveness, any translation or implementation of MAC in the context of a social, economic, or political prisoner's dilemma could have undesirable repercussions.

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.

Axelrod, R. and W. Hamilton. 1981. The evolution of cooperation. *Science* 211:1390-96.

Axelrod, R., and D. Dion. 1988. The further evolution of cooperation. *Science* 242:1385-90.

Boyd, R., and J. Lorberbaum. 1987. No pure strategy is evolutionarily stable in the repeated prisoner's dilemma game. *Nature* 327:58-59.

Friedman, J. 1971. A non-cooperative equilibrium for supergames. *Review of Economic Studies* 38:1-12.

Harris, R. 1969. Note on "optimal" policies for the prisoner's dilemma. *Psychological Review* 76:373-75.

Luce, R., and H. Raiffa. 1957. *Games and decisions.* New York: Wiley.

Marinoff, L. 1990. The inapplicability of evolutionarily stable strategy to the prisoner's dilemma. *The British Journal for the Philosophy of Science* 41:461-72.

Maynard Smith, J. 1982. *Evolution and the theory of games.* Cambridge: Cambridge University Press.

Weatherford, R. 1982. *Philosophical foundations of probability theory.* London: Routledge & Kegan Paul.