Centrum voor Wiskunde en Informatica

REPORT*RAPPORT*

The Influence of Evolutionary Selection Schemes on the Iterated Prisoner's Dilemma

D.D.B. van Bragt, C.H.M. van Kemenade, J.A. La Poutré

Software Engineering (SEN)

SEN-R0004 February 28, 2000

Report SEN-R0004
ISSN 1386-369X

CWI
P.O. Box 94079
1090 GB Amsterdam
The Netherlands

The Influence of Evolutionary Selection Schemes on the Iterated Prisoner's Dilemma

D.D.B. van Bragt, C.H.M. van Kemenade, and J.A. La Poutré
CWI
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands
{bragt,kemenade,hlp}@cwi.nl

ABSTRACT

Many economic and social systems are essentially large multi-agent systems. By means of computational modeling, the complicated behavior of such systems can be investigated. Modeling a multi-agent system as an evolutionary agent system, several important choices have to be made for evolutionary operators. Especially, it is to be expected that evolutionary dynamics substantially depend on the selection scheme. We therefore investigate the influence of evolutionary selection mechanisms on a fundamental problem: the iterated prisoner's dilemma (IPD), an elegant model for the emergence of cooperation in a multi-agent system.

We observe various types of behavior, cooperation level, and stability, depending on the selection mechanism and the selection intensity. Hence, our results are important for (1) The proper choice and application of election schemes when modeling real economic situations and (2) assessing the validity of the conclusions drawn from computer experiments with these models. We also conclude that the role of selection in the evolution of multi-agent systems should be investigated further, for instance using more detailed and complex agent interaction models.

1991 Mathematics Subject Classification: 68T05, 68T20, 90D80
1991 ACM Computing Classification System: G.1.6, I.2.6, I.2.8
Keywords and Phrases: agent-based economics, evolutionary algorithms, evolutionary selection schemes, iterated prisoner's dilemma
Note: Work carried out under theme SEN4 "Evolutionary Systems and Applied Algorithmics".
This paper will appear in the Journal of Computational Economics and has been presented at the 5^{th} Int. Conf. on Computing in Economics and Finance (CEF'99) (Boston College, Chestnut Hill, MA, June 24-26, 1999), the 11^{th} Belgium-Netherlands Conference on Artificial Intelligence (BNAIC'99) (Maastricht, the Netherlands, November 3-4, 1999), and the 9^{th} Belgian-Dutch Conf. on Machine Learning (BENELEARN'99) (Maastricht, the Netherlands, November 5, 1999).
The authors would like to thank all conference participants for their feedback. Special thanks are also due to Bill Langdon, Bruno Beaufils, Tomas Klos, and Dirk Thierens for their constructive comments and useful suggestions.

1. Introduction

Recently, interest in the evolution of strategic choice and the simulation of adaptive agents has surged among economists and game theorists [2]. We simulate the dynamic behavior of multi-agent systems with evolutionary algorithms (EAs). EAs are now regarded as a powerful means for modeling social learning processes occurring in economic multi-agent systems [7]. In an evolutionary context, learning takes place in three different steps: (1) learning by imitation (reproduction and selection), (2) communication and exchange of strategic information (recombining or "crossing over" genetic information), and (3) experimentation (mutation of the strategies). In this paper we focus on the modeling of agent selection in EAs. An economic interpretation of the considered selection schemes is given as well, connecting the fields of evolutionary computing and agent-based economics.

Our simulations show that the course of evolution in a multi-agent system is sensitive to the selection intensity, whether the best agents are retained in future generations, and the stochastic nature of evolutionary processes. These, in turn, indicate the potential sensitivity of the dynamics and, quite importantly, the level of cooperation in multi-agent systems to the selection of the individual agents.

The interaction among agents in their ecological "societies" is modeled here with the well-known iterated prisoner's dilemma (IPD) game. The IPD game has been studied intensively in the past (see Section 2), especially to investigate the emergence of cooperation in a society of selfish agents. We extend previous work by investigating the influence of evolutionary selection schemes on the dynamics of a population of agents playing the IPD game. In computer science, many evolutionary selection schemes have already been proposed and evaluated in the field of numerical optimization [6]. This field is chiefly concerned with generating optimal solutions for mathematical optimization problems (in which the fitness of the evolving agents is determined by a *fixed* objective function). We, by contrast, focus on processes in which the fitness of the agents depends on the strategies of the other agents in the population. These so-called "ecological" problems are of interest, because qualitative aspects of population dynamics (e.g., the occurrence of transients or the emergence of stable long-term behavior) become very important.

This paper is organized as follows. First, the prisoner's dilemma is introduced in Section 2. A brief description of our computational experiments is given in Section 3. Section 4 provides an overview and discussion of the main results. First, pure selection schemes, in which all parents are replaced by offspring in the next generation, are analyzed. Then, elitist models, in which the best agents remain present in the next generation, are evaluated. Both mean population dynamics and results from individual experiments are included in these evaluations. Conclusions are given in Section 5.

2. The prisoner's dilemma

The basic prisoner's dilemma (PD) is a two-player game. Both players can choose between two options: "cooperate" or "defect". If both players cooperate, they receive a reward payoff R; if both defect, they receive a punishment P; if one player cooperates and the other defects, the cooperator receives the sucker's payoff S and the defector the temptation payoff T. To constitute a PD game, the payoffs should satisfy $T > R > P > S$ and $2R > S + T$. The payoffs used in the current experiments, as in most other IPD studies, are $T = 5$, $R = 3$, $P = 1$ and $S = 0$ [1, 10, 4, 3, 5].

An interesting variation of the single-round prisoner's dilemma is the iterated PD (IPD), in which two players repeatedly play a PD game against each other. Axelrod [1] is the first to demonstrate, using a genetic algorithm (GA), that it is possible to simulate the emergence of robust cooperative behavior in a group of interacting IPD players. Miller [5] uses a GA to perform similar experiments with co-evolving finite state machines (FSMs). Fogel [3] evolves a population of FSMs with evolutionary programming techniques. Other important contributions include Lindgren's [4] ecological experiments with adaptive-memory players and Yao and Darwen's [10] co-evolutionary simulations of the N-person IPD. However, a rigorous investigation of the influence of selection schemes on the evolutionary IPD has not been performed before.

We show that the particular scheme governing the agents' selection has a strong influence on the emergence of cooperation and the stability of the evolving populations. Hence, our work extends previous results and leads to further insight into the dynamics of multi-agent systems.

3. Experimental setup

The computer experiments are performed using evolutionary algorithms. These algorithms are stochastic search methods based upon the principles of natural genetic systems [6]. EAs deal with a population of individuals (*agents*). In this paper, simulations are performed with a population of 60 agents. In each fitness evaluation, an agent competes against 12 randomly selected (without replacement) opponents. The fitness of an agent is defined as its mean payoff over all played games. To reduce the computational demands, the IPD game is terminated after 20 rounds.

The IPD strategy of each agent is coded as a bitstring (called a "chromosome") of finite length. We use the genetic representation of IPD players proposed in [1]. Each agent in our model has a memory capacity of four previous moves (two of his own and two of the opponent). The agent's strategy, which specifies the agent's next move, is then encoded in $2^4 = 16$ bits, located on the agent's chromosome. Four additional bits are present to initialize the agent's memory of past moves at the start of play. Hence, the total chromosome length is 20.

The EA begins with a randomly generated population. Subsequent populations of the same size are generated using (1) selection, (2) crossover, and (3) mutation. The different selection schemes studied in this paper are described in detail in Section 4. After the selection phase, each pair of "parent" agents produces two "offspring", whose chromosomes are determined by crossing over the parental chromosomes. Finally, the chromosomes are mutated with a certain (small) probability P_m. Here, the mutation probability P_m is set to a value of 0.025 per bit.

4. Results and discussion

Evolutionary selection schemes can be characterized as either generational or steady-state schemes [8]. We focus on the commonly used generational schemes, in which each population is replaced in one step by a new population. In the less frequently used steady-state models, only a small number of agents is replaced each generation.

A further subdivision can be made between pure and elitist selection schemes. Pure selection schemes (see Section 4.1) allow no overlap between successive generations: all parents from generation G_t are discarded and the next generation G_{t+1} is filled entirely with offspring from these parents. In elitist schemes, subsequent generations typically do overlap (see Section 4.2): well-performing parents are transferred to the next generation and only the badly-performing parents are replaced.

We do not scale our results with respect to the number of function evaluations. This would be an appropriate procedure in optimization work (where computational efficiency is very important). However, for the ecological processes we study here, the computational effort is of secondary importance.

4.1 Pure selection schemes

The following selection schemes are considered in this section:

- Canonical genetic algorithms (CGAs). These traditional GA algorithms use fitness-proportional selection of the parents in combination with single-point crossover [6]. The crossover probability P_c is set to a commonly used value of 0.6. In addition, comparative studies with other recombination models (e.g., with multiple crossover points, or with a higher crossover probability) have been performed. The results do not change significantly, however, indicating that the findings presented in this paper are not very sensitive to this modeling element.

- Tournament selection GAs (TGAs). This is an example of a local selection method. A random uniform sample of size $q > 1$ is taken from the population [6]. From this sample, only the fittest agent is selected to produce offspring. The selection intensity becomes larger if the tournament size q increases.

We first consider various CGA variants. Figure 1 shows that no clear population dynamics emerge if agents are selected to reproduce in proportion to their absolute fitness in the current generation (the CGAa scheme). This global selection method exercises a rather weak selection intensity. The mean fitness of the agents remains close to 2 (in between the punishment payoff P=1 and the reward payoff R=3), i.e., the population is neither evolving to a society of defectors (with a mean fitness close to P) or a society of cooperating agents (with a mean fitness close to R).

Axelrod [1] and Miller [5] use a variant of the CGA model called "sigma scaling" [6], in which the agent's raw fitness f_i is normalized by taking $\hat{f}_i = (f_i - \mu)/\sigma + \alpha$, where μ is the population mean fitness and α determines the influence of the agent's relative performance. Negative fitness values ($\hat{f}_i < 0$) are set equal to zero, which means that agents with a fitness more than α standard deviations below the mean get no offspring.

It is important to note several important properties of sigma scaling. First, this selection method is immune to (strictly increasing) affine transformations of the payoff matrix. This feature is attractive, because players should be able to recalibrate their utility scales without consequences. Second, problems with negative (raw) fitness values (in case of negative payoff values) do not occur using sigma scaling because this method only considers the relative performance of agents. Third, the selection intensity remains constant during the run because each agent's fitness is scaled with respect to the fitness variances in the population.

Results for two different values of α are shown in Fig. 1. The $\alpha = 1$ model (used in [1]) is labeled as CGAb, and the $\alpha = 2$ case (see [5]) is labeled as CGAc. Notice the emergence of

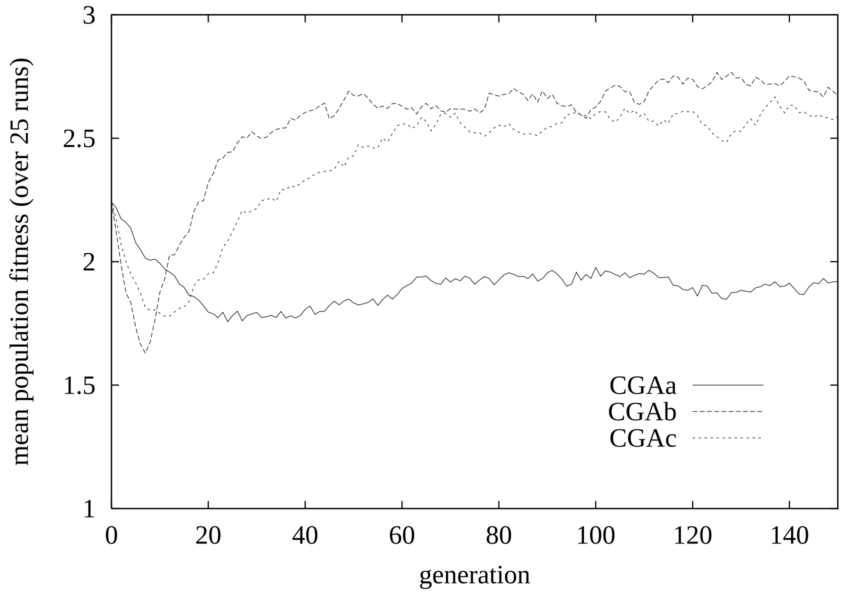

Figure 1: Evolution of the mean population fitness using (1) the canonical GA (CGAa) and (2) sigma scaling of the fitness values with $\alpha = 1$ (CGAb) and $\alpha = 2$ (CGAc). Notice the emergence of distinct cooperative behavior if agents are not selected in proportion to their *absolute* performance (as in the CGAa model) but in proportion to their *relative* performance (as in the CGAb and CGAc models).

a clear population dynamic, which contrasts with results using the previous (CGAa) model. The mean fitness of the population initially decreases toward the punishment level P (=1), indicating that defecting agents are spreading throughout the population. After approximately 10 generations, some agents begin reciprocating cooperation (while still punishing a defecting opponent). The fitness of these cooperative agents is significantly higher than P, because they receive the reward payoff R (=3) whenever they meet each other. Hence, cooperative agents proliferate and the mean population fitness increases.

An important conclusion at this point is that cooperation does not emerge in the IPD if agents are selected in proportion to their absolute performance, as is done in a classical selection scheme. Cooperation is established, however, if selection is based upon the relative performance of the agents. Competition between agents is situated at a global level in these canonical GAs, because an agent's performance depends on the performance of all other agents as well.

In the remainder of this section we consider a different class of selection schemes: tournament GAs. This class of selection schemes, in an economic context, is an appropriate model for a more local selection of agents.

Figure 2 shows results for the TGA scheme for different tournament sizes q. Comparing Fig. 2 with Fig. 1 we see that the TGA scheme with $q = 2$ yields a population dynamic similar to the fitness-proportional selection schemes based upon relative performance (CGAb and CGAc). The TGA scheme allows a detailed study of the influence of the selection intensity by changing the tournament size q. The selection intensity can be interpreted in

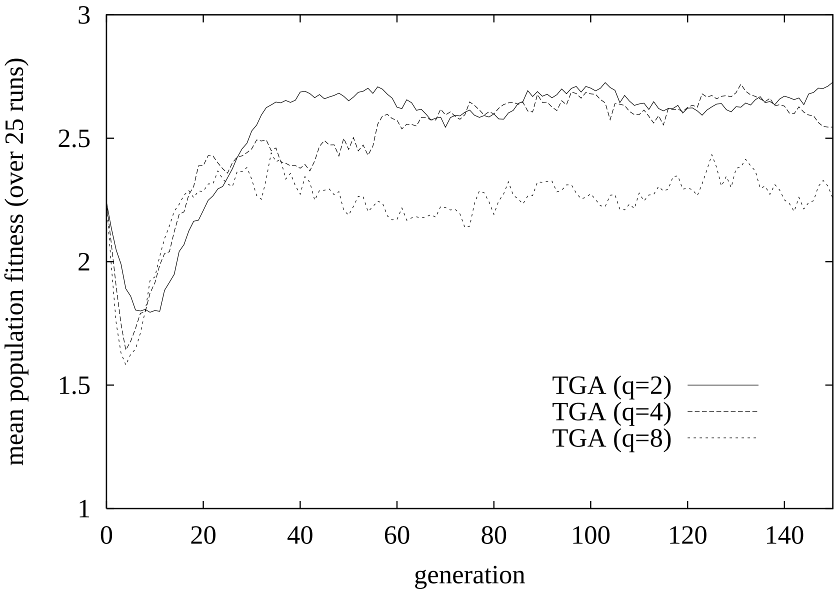

Figure 2: Evolution of the mean population fitness using tournament selection (TGA) with different tournament sizes ($q = 2$, $q = 4$, and $q = 8$). Notice the decrease in long-term fitness if the selection intensity (i.e., the level of competition between agents) becomes very large, e.g., for q equal to 8.

an economic context as a parameter determining the level of competition between different agents. Increasing the selection intensity by increasing the tournament size q, the mean long-term fitness decreases, see Fig. 2. This implies that, if the selection intensity becomes very large, populations of agents with high fitness (i.e., a high level of cooperation) are being exploited more successfully by non-cooperative agents.

The results shown in Fig. 2, averaged over 25 experiments, mask a rather unpredictable population dynamic in individual experiments. As an example, Fig. 3 shows results for five individual runs for the TGA selection model (with $q = 2$). Figure 3 shows that the evolution of a population strongly depends on the stochastic nature of the evolutionary processes. Notice, for instance, that, from a random start, the population sometimes converges very rapidly to a cooperative state (within 20 generations), while in other experiments mutual cooperation is reached only after 50 generations or more. Once a cooperative society has been established, it remains vulnerable to invasions of exploiting (non-cooperative) agents - notice the occasional collapses of the population fitness in Fig. 3.

Additional simulations show that the transitions between the two extremal states P and R occur more frequently if the tournament size q increases. An example is given in Fig. 4, where the tournament size is increased to $q = 8$. The increased selection intensity clearly leads to very unstable behavior of the evolving populations. This implies that, in an economic system, cooperation between groups of agents is most likely to be stable if the level of competition does not exceed certain thresholds. But, we observed earlier in the CGAa model that cooperation does not emerge at all if competition is too weak. Hence, we conclude that stable cooperation between agents is most likely to emerge in multi-agent systems with an intermediate level of

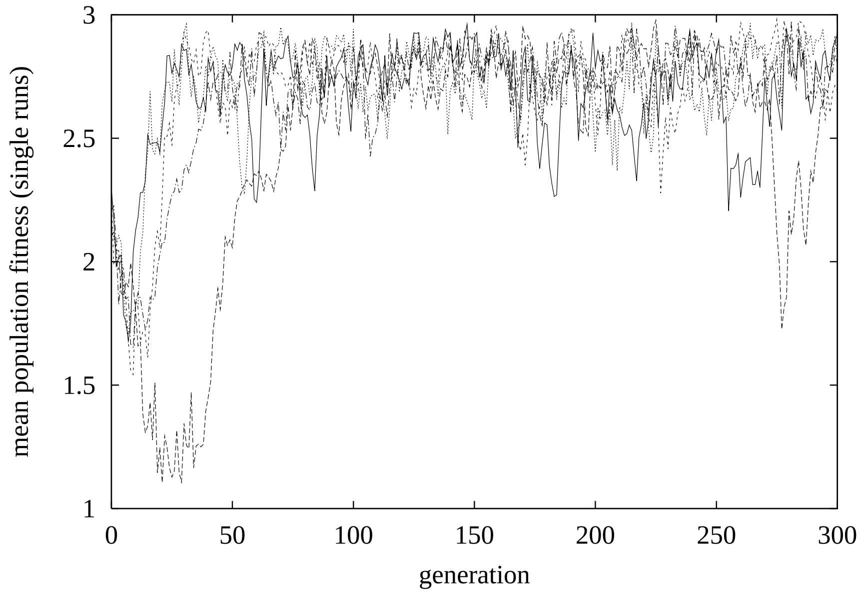

Figure 3: Evolution of the mean population fitness in five individual experiments, using the TGA selection scheme (with $q = 2$).

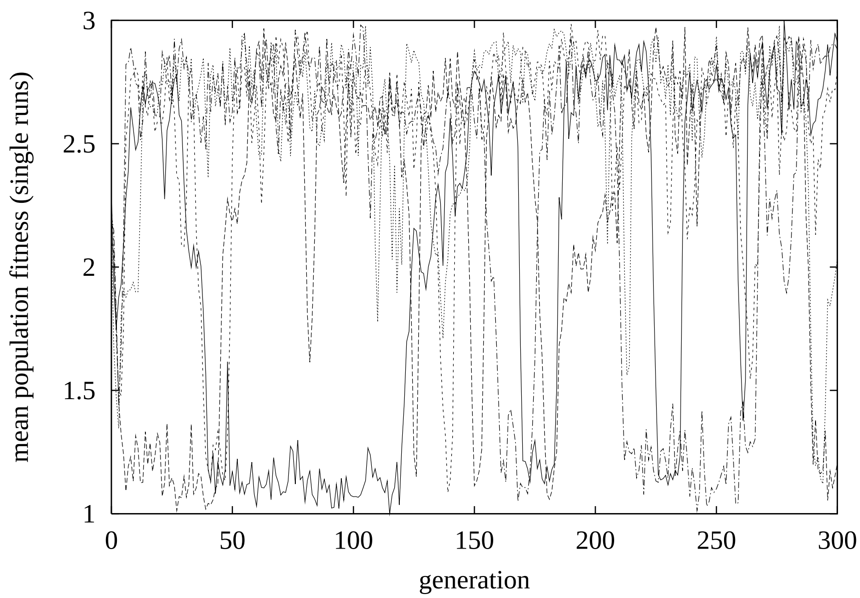

Figure 4: Evolution of the mean population fitness if the tournament size q increases to 8 in the TGA selection scheme. Notice the unstable population dynamics in comparison with Fig. 3.

competition.

It could be conjectured that a loss of diversity in agent types causes the significant decrease of evolutionary stability if the selection intensity increases. Experiments with the TGA

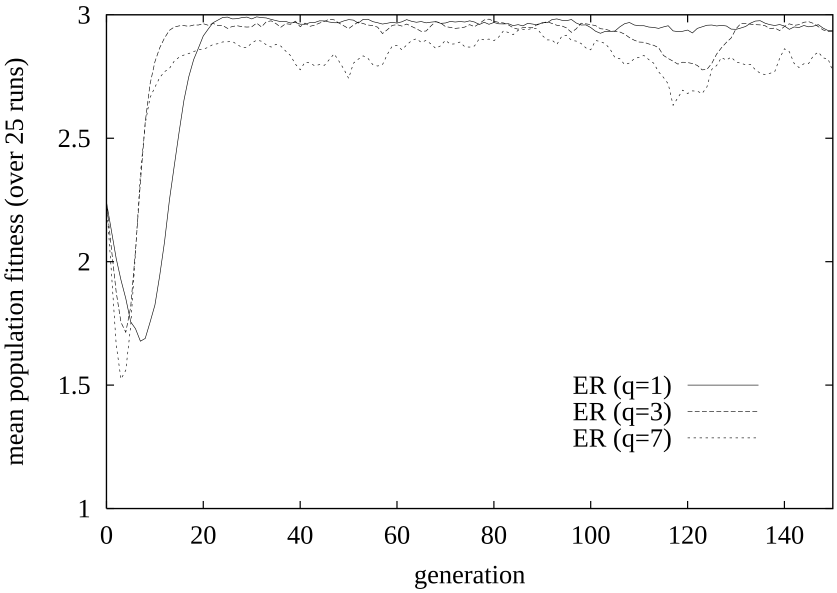

Figure 5: Evolution of the mean population fitness for the elitist recombination (ER) scheme with different tournament sizes ($q = 1$, $q = 3$, and $q = 7$).

scheme (with $q = 8$) show, however, that the number of genotypes present in the population remains high (on average above 50). Considering that the population size (and hence the maximum number of different genotypes) is equal to 60, it becomes plausible that a loss of diversity is not the main reason for the observed decrease in evolutionary stability. Indeed, we doubled the population size from 60 to 120 to increase the population diversity even further, and again an increasing selection intensity was accompanied by a decreasing evolutionary stability, as in the experiments with a smaller population size.

4.2 Elitist selection schemes

As an important instance of an elitist model, the elitist recombination (ER) scheme [9] is considered. In this scheme, the population is first shuffled randomly and partitioned in pairs of parents. Then, each pair of parents creates two offspring, and a local competition between parents and their offspring is held. Finally, the best two agents of each "family" are transferred to the next generation. This might be a suitable model for economic selection at the "family" or group level. In this case, new agents combine strategies of (a pair of) successful agents in an attempt to outperform them and gain access to the market.

Given the importance attributed to the selection intensity in the previous section, it is natural to investigate the same issue for elitist schemes. Increasing the selection intensity of elitist recombination is discussed in [8], where it is proposed to adjust the selection intensity in the ER scheme by selecting one of the parents in a tournament of size q (whereas the other parent is still selected at random).

Figure 5 shows the evolution of the mean population fitness for this selection scheme for three different tournament sizes q. Previous work [8] shows that the selection intensity of the

ER ($q-1$) scheme is equal to the selection intensity of the TGA (q) scheme.[1] We therefore present results for tournament sizes q of 1, 3, and 7 in Fig. 5. Comparing the results of the TGA ($q=2$) scheme in Fig. 2 with those of the ER ($q=1$) scheme, it is clear that an elitist scheme yields a society of agents with a much higher cooperation level. The long-term fitness decreases somewhat with increasing q in Fig. 5. This effect is, however, relatively small for this elitist scheme in comparison with pure schemes (see Fig. 2).

Additional experiments show that the sensitivity to stochastic processes decreases if the agents are selected using an elitist instead of a pure model. In particular, the collapses in fitness, as observed in Figure 3, occur less frequently in individual experiments with the ER ($q=1$) scheme, see Fig. 6. This indicates that emerging cooperative societies are less

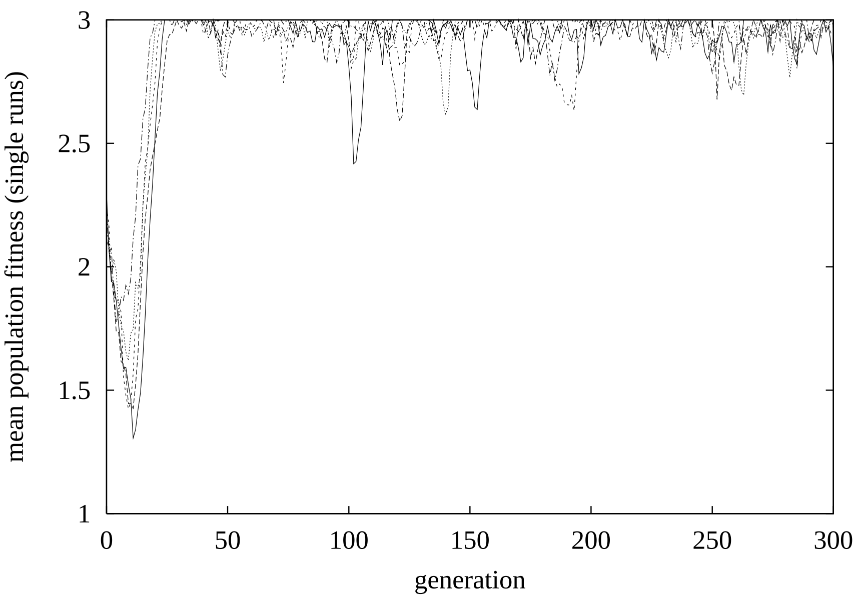

Figure 6: Evolution of the mean population fitness for the elitist recombination scheme (ER) with a tournament size q equal to one. Notice the fast convergence of the evolving populations towards a very high level of cooperation.

vulnerable to exploitation by defectors if an elitist model governs agent selection.

We have shown in this section that highly-cooperative societies can evolve if the best performing agents are always retained in the population ("elitism"). Elitism also has an important stabilizing effect: the mean performance of the agents remains high even if competition becomes intense. This demonstrates that economic systems with a certain degree of conservatism are less vulnerable to unpredictable behavior and low performance. Also, the elitist recombination model might be a suitable model for economic selection at the group or "family" level. In this case a direct competition exists between well-performing "parent" agents and new "offspring" agents (trying to enter the market by combining the parental strategies in an innovative manner).

[1] Assuming that the heritability factor h^2 is equal to one (for details, see [8]).

5. Conclusions

We study evolutionary processes in economic multi-agent systems. As an example, we investigate the "evolution of cooperation" in a population of agents playing the (iterated) prisoner's dilemma (IPD). Computational experiments are performed using evolutionary algorithms (EAs). Various selection schemes proposed in this field are evaluated and an economic interpretation of these schemes is given.

Major differences in population dynamics are observed, depending on the agent selection mechanism and the level of competition between the agents. In systems without a form of conservatism (or "elitism"), robust cooperative societies emerge only if the level of competition between the agents is neither too small nor too large. If agent selection is elitist (when the best agents remain in the population) stable societies of highly-cooperative agents evolve. These results strongly indicate that a precise modeling of selection processes is of crucial importance when modeling economic situations. The validity of computer experiments thus critically depends on a correct mapping of real economic processes to more abstract simulations.

The results presented in this paper are but a first step in achieving a solid understanding of the complicated, but fascinating, dynamics of multi-agent systems. Much work has yet to be done using more detailed agent interaction models.

References

1. R. Axelrod. The evolution of strategies in the iterated prisoner's dilemma. In Lawrence Davis, editor, *Genetic Algorithms and Simulated Annealing*, pages 32–41. Pitman, London, 1987.
2. K. Binmore and N. Vulkan. Applying game theory to automated negotiation. *Netnomics*, Vol. 1(1):1–9, 1999.
3. D.B. Fogel. Evolving behaviors in the iterated prisoner's dilemma. *Evolutionary Computation*, 1(1):77–97, 1993.
4. K. Lindgren. Evolutionary phenomena in simple dynamics. In C.G. Langton, C. Taylor, J.D. Farmer, and S. Rasmussen, editors, *Artificial Life II, Vol. X*, pages 295–312. Addison-Wesley, 1991.
5. J.H. Miller. The coevolution of automata in the repeated prisoner's dilemma. *Journal of Economic Behavior and Organization*, 29(1):87–112, 1996.
6. M. Mitchell. *An Introduction to Genetic Algorithms.* MIT Press, 1996.
7. T. Riechmann. Learning and behavioral stability - An economic interpretation of genetic algorithms. *Journal of Evolutionary Economics*, 9(2):225–242, 1999.
8. D. Thierens. Selection schemes, elitist recombination, and selection intensity. In Th. Bäck, editor, *Proceedings of the 7th International Conference on Genetic Algorithms*, pages 152–159. Morgan Kaufmann, 1997.
9. D. Thierens and D.E. Goldberg. Elitist recombination: An integrated selection recombination GA. In *Proceedings of the First IEEE Conference on Evolutionary Computation*, pages 508–512. IEEE Press, 1994.
10. X. Yao and P.J. Darwen. An experimental study of N-person iterated prisoner's dilemma games. *Informatica*, 18:435–450, 1994.