Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

Learning, Evolution and Population Dynamics

by

Jürgen Jost, and Wei Li

World Scientific
www.worldscientific.com

LEARNING, EVOLUTION AND POPULATION DYNAMICS

JÜRGEN JOST and WEI LI*

*Max Planck Institute for Mathematics in the Sciences,
Inselstr. 22, 04103 Leipzig, Germany*

We study a complementarity game as a systematic tool for the investigation of the interplay between individual optimization and population effects and for the comparison of different strategy and learning schemes. The game randomly pairs players from opposite populations. It is symmetric at the individual level, but has many equilibria that are more or less favorable to the members of the two populations. Which of these equilibria is then attained is decided by the dynamics at the population level. Players play repeatedly, but in each round with a new opponent. They can learn from their previous encounters and translate this into their actions in the present round on the basis of strategic schemes. The schemes can be quite simple, or very elaborate. We can then break the symmetry in the game and give the members of the two populations access to different strategy spaces. Typically, simpler strategy types have an advantage because they tend to go more quickly toward a favorable equilibrium which, once reached, the other population is forced to accept. Also, populations with bolder individuals that may not fare so well at the level of individual performance may obtain an advantage toward ones with more timid players. By checking the effects of parameters such as the generation length or the mutation rate, we are able to compare the relative contributions of individual learning and evolutionary adaptations.

Keywords: Evolutionary complementarity game; individual learning; population dynamics; evolutionary adaptation.

1. Introduction

Our first aim is to investigate the relation between individual optimization and the resulting collective dynamics in an evolving environment. This topic has a long history, starting (at least) with Mandeville's essay [12] on bee colonies and Adam Smith's [22] invisible hand. From a scientific point of view, it is important to analyze the validity and generality of the many claims that have been put forward since then and to identify their necessary and sufficient assumptions. For that purpose, we need a simplified formal model in which one can isolate the key mechanisms and features without all the contingent details of real world situations. In order to proceed in that direction, we utilize an agent-based model that can be readily simulated and

---

*Current address: Santa Fe Institute, E-mail: liw@santafe.edu.

is also amenable to formal analysis. This model is a population game where the equilibrium at the individual level is degenerate so that the selection among the possible equilibria results from the collective dynamics at the population level.

Our second aim is to compare the strength of different learning schemes in an evolving competitive situation, i.e. where the opponents also try to learn efficiently. For this purpose, in the sense of statistical learning theory, every agent needs to have a stream of stochastic input data on the basis of which he can develop his models. In order to overcome the limitations of classical game theory, an agent encounters in each round an opponent that is randomly chosen from an opponent population. When all agents in one population employ a particular learning strategy, and all agents from the opponent population employ another strategy, we can see in favor of which population the equilibrium develops.

Our third aim is to connect and compare the two previous aspects, evolution and learning. Thus, we want to see what is better for agents and for populations of agents: to adapt by evolution or to learn by individual experience.

Our model is the following complementary game played between members of opposite populations, as introduced in Ref. 10: A buyer and a seller meet and independently each makes an offer between 0 and K (K is a sufficiently large integer, usually taken to be 50 in our simulations). If the buyer's offer k_b is at least as large as the seller's offer k_s, a deal is concluded and the buyer gains $K - k_b$, the seller k_s; otherwise, they gain nothing. Thus, in order to be the most successful, the buyer should offer not less than what the seller is asking, because otherwise he will not get a deal, but also not much more, in order not to pay too much. We note that every integer between 0 and K is an equilibrium in the sense that no player can do any better than playing that value if his opponent does so. This game is then played repeatedly between members of two opposing populations of the same size: the buyers and the sellers. In each round, the members of the two populations are randomly paired, i.e. every buyer is paired with a randomly chosen seller. After a fixed number of rounds, the accumulated gains of the agents in each population are compared, and on the basis of this fitness function some evolutionary scheme constructs a new population. Thus, the basic situation is symmetric between the two players, the buyer and the seller, and also between the two populations. We can then break that symmetry by equipping the buyers and the sellers with different strategy spaces. The differences here could simply be differences of memory span, i.e. how many encounters an agent can remember and utilize for determining his own current bid. The agents could also employ totally different strategies. Here, the possible strategies could range from playing a random number to an elaborate scheme for computing a bid on the basis of all information available from the agent's experience or even including the experiences of his friends in his own population.

In general, the agents of each population will adapt through individual learning and through fitness-based evolution. Thus, even if the agents in a population act completely independently of each other, they will feel the long term effects of the

actions of their fellows through the collective adaptation of the agents of the other populations.

If the members of one population — say, the sellers — could coordinate their actions, their best strategy would consist in always choosing that bid that is optimal for them, K in this case. The other population would then have no choice but to accept that and also play the same bid. However, as long as the buyer population has not evolved to that state that is so unfavorable to them, it pays for any individual seller to lower his bid and to increase his chance for a successful deal. Thus, such an agent would be more successful than the ones keeping to the population optimum K, and because of his higher fitness, his strategy would then be more frequently represented in the next generation. Thus, in this evolving scheme, the group optimum is not stable against defections in the own population. In particular, those agents in a population that are individually the fittest can cause a decrease in the fitness of the population as a whole. This relates to our first aim.

Also, since the members of both populations are trying to maximize their fitness, one cannot expect that either population can enforce that equilibrium that is optimal for itself on the other population. In a symmetric situation, we would expect $K/2$ to be the eventual steady state. When the buyers and the sellers employ different strategies, we can simply decide which strategy is superior by checking whether the steady state reached in that situation is smaller or larger than $K/2$. In the first case, the buyers are doing better; in the second, the sellers. This then allows us to address our second aim.

We can also play with such parameters as the generation length or the mutation rate in the evolutionary step. In that way, we can compare the relative contributions of individual learning and evolutionary adaptations, as formulated in our third aim.

Although the rules of our game are extremely simple, the action takes place at three different levels: the individual agents evaluate the information they obtain from their interactions and use that to compute their next own actions (first level — information evaluation and learning), they compete with each other inside a population (second level — adaptation and evolution), and the populations are compared with each other (third level — competition between strategy spaces). The link between the first two levels is provided by the fitness function; the one between the higher levels, by the collective dynamics at the population level resulting from the individual optimizations.

While what we present here clearly does not yet constitute a complete theory, we believe that we have found a formal model that on one hand is simple enough for easy simulations and also permits formal analysis, but on the other hand is rich enough to capture many of the essential features of the conceptual problems we wish to address.

We shall first put our model in the perspectives of game theory and of certain traditions in economics. After that, we shall start with some mathematical considerations before we present various simulations that both illustrate some of our formal reasoning and yield insightful results beyond those that we can demonstrate formally.

2. Game Theory

In order to put the present work into perspective, we now shall discuss in more detail how it fits into modern game theory. In the classical model of game theory, as introduced by von Neumann and Morgenstern [27], we have two players that meet once. They have a finite set of action options, and the payoff for each of them is determined by his own and his opponent's action. They are both perfectly rational and possess and can utilize all relevant information. Thus, they both try to maximize their payoff function, each also knowing his opponent's payoff function and therefore choosing that action that best anticipates the opponent's move, which is assumed to be anticipating in the same way. In such a situation, there exists a Nash equilibrium [15] in which no player can improve his payoff by changing his action, given that the opponent will react correspondingly in his own best interest. That Nash equilibrium need not be unique; in fact, in our basic model game, each value k between 0 and K, when played by both players, is a pure strategy Nash equilibrium. In addition, there are mixed strategy equilibria. In particular, there is no rational way for a player to decide which bid to play because each bid is a pure Nash equilibrium when also played by his opponent.

We are therefore interested in the mechanisms that can select between all those equilibria. Since in our model the game is played repeatedly, and the agents can benefit from their own experience (or those of other players in some version of our game), this brings us to the theory of learning in games; see e.g. Ref. 5 as a reference for our discussion of this topic. Also, the game is played in populations of agents, which leads us into evolutionary game theory, where we can use, for example Ref. 28 as a reference. Leaving the issue aside for the moment that in our model we have two distinct populations, we are considering a random-matching model where in each round all players are randomly matched, but can observe only their own matches. Also, in our model, no player acts in the interest of the population, but only myopically pursues his own aims. With regard to learning and evolution, we have a case of fictitious play where players only observe (and perhaps memorize) the results of their own matches. (Nevertheless, we shall also consider scenarios where players have information about the performance of selected other players in their own population — their "friends.") The players then evolve after many rounds of play according to their relative fitness in their population (as in evolutionary algorithms), in contrast for example to replicator dynamics, where the relative frequencies of strategies continuously change according to their actual performance [9]. In fact, we shall discuss below the issue of generation length, i.e. after how many rounds the fitness of the players is evaluated; the extreme case of generation length 1 could be made to correspond to discrete time replicator type dynamics. Such an on-line evolutionary adaptation, however, would prevent individual agents from improving their performance on the basis of their own experience, i.e. learning. In other words, we are interested in a hybrid of games with learning and evolutionary games, in order to assess the relative strengths and problems of the two schemes.

Coming to evolutionary aspects, the concept of an evolutionarily stable strategy (ESS) as introduced by Price and Maynard Smith [13] is not directly applicable to our setting because the matches are played between members of different populations. (Also, as emphasized for instance in Ref. 26, the concept of an ESS takes as its base situation a monomorphic population, i.e. one where all members utilize the same strategy. In many applications, however, one is naturally interested in the stability of a polymorphic population against invasions of mutants. But, in our case, the basic equilibria that we shall take as our default situation do consist of monomorphic populations applying pure strategies.) Versions of evolutionary stability for multipopulation games have been developed in Refs. 24 and 2, for instance (see Ref. 28 for further references). In those definitions, a population with a rare mutant in one of the populations is compared with the original population in its performance against the other unchanged populations. By the result of Ref. 20, a strategy is evolutionarily stable iff it is a strict Nash equilibrium. In particular, when the game is at a pure Nash equilibrium, any mutant strategy performs less well than the dominant strategy — *provided that the opposite population does not adapt to those mutants.* This provision shows that the concept of evolutionary stability is essentially a static one and therefore not so well suited for our setting. Whether it will be advantageous for members of a population to adapt to a rare mutant in the opposite population, however, depends on the relations between the various parameters, as the following heuristic reasoning shows. If the generation length is short, i.e. each player plays only a few rounds before his fitness is evaluated, then it will not pay off to adapt to a possible rare mutant in the opposite camp. Namely, in that situation, a player is unlikely to encounter such a rare mutant in his lifetime, and there is this point in playing a strategy that is most likely to be inferior in all encounters. If the number of matches played becomes higher, comparable to the size of the population, then there is a substantial chance to encounter a rare mutant at least once. Let us assume that the player is a buyer, and the rare mutant is a seller that asks a higher price — say, $k+1$ — than the population equilibrium k. If the buyer offers $k+1$ in all m matches, his accumulated gain is $m(K-k-1)$. If he offers only k and encounters the mutant once, he will accumulate $(m-1)(K-k)$. So, adapting to the mutant will be advantageous when $m < K - k$. As a typical evolutionary outcome of our game is that $k \sim K/2$, this means that evolutionary stability will depend on the population size. In the preceding heuristics, we have assumed that there is only a single mutant. If there are more of them, of course, they can destabilize the reaction of the opposite population more easily. In any case, we arrive at the heuristic conclusion that in a sufficiently large population, invasions of rare mutants should not lead to a response adaptation of the other population that ultimately makes the performance of the mutants superior. This is, of course, in accordance with the general idea of an ESS. The issue of stability of strategies in multipopulation games can, however, also be addressed through the inherently dynamical framework of replicator dynamics. This approach has been developed in Refs. 1, 8 and 28. In particular, in our game, mixed Nash equilibria

are no longer (Lyapunov or asymptotically) stable for the replicator dynamics, as follows from the analysis of Ref. 10. The pure Nash equilibria, however, remain stable attractors, and which one is ultimately approached depends on the initial conditions. This fits, of course, the general theory. A strategy combination of the two populations is evolutionarily stable iff it is a strict Nash equilibrium (which necessarily is pure) (by Ref. 20, as already mentioned) iff it is asymptotically stable for the standard replicator dynamics [7, 8, 28]. In a similar direction, we have the regular and payoff monotonic evolutionary systems of Ref. 26, a replicator type dynamical model of multipopulation games, whose asymptotically stable states also yield Nash equilibria.

In summary, in the context of game theory, the distinctive features of our model are the following. *A priori* rational reasoning, as in the standard game-theoretic paradigm, does not lead to a unique solution because there exist multiple equilibria. Which one is achieved depends on the historical contingencies of the evolution process. Those in turn are governed by the available strategic options for utilizing the information obtained by repeated interactions with different opponents. Since we have a two-population game, we can equip the two populations with different strategy spaces, particularly different learning schemes, and can then see which one fares better. In other words, the game-theoretic degeneracy of our model allows us the comparison of evolutionary and learning mechanisms that can break that degeneracy favorably. In our game, players can have full information about the payoffs of themselves and their opponents, but this information is not so helpful as the population dynamics are not only shaped by their own actions and the responses of their opponents, but also by those of their conspecifics with whom they compete via their fitness function. The players, however, do not have information about the previous actions of their current opponents, but only information about the actions of their previous opponents in their past encounters with them. We shall also investigate the situation where a player also has information about the behavior of the opponents of some other members of his own population, his "friends," but since in any case the opponents are randomly sampled, this will not lead to decisive advantages in our evolutionary simulations. Thus, for instance, some of the learning schemes analyzed in Ref. 26 are not applicable. In particular, a player cannot distinguish between his different opponents and so there is no way a player could respond to his opponent individually.

3. Economic Ideas About the Evolution of Institutions

While we have mentioned Adam Smith in the introduction, a conceptually closer starting point for our model is the work of Carl Menger [14], the founder of the Austrian school of economics, with his (direct or indirect) disciples F. von Wieser, E. von Böhm-Bawerk, L. von Mises, F. von Hayek and J. Schumpeter (see e.g. Ref. 19). For Menger, economic phenomena were the emergent product of individual rational actions in situations where information is meagre and where actions are costly.

Methodological individualism kept Menger and his school critical of techniques of aggregations as they underly neoclassical economics (see e.g. the standard reference Ref. 17), and of ideas of central planning or welfare economics, and emphasis on rational actions clearly distinguished them from philosophies like utilitarianism. More recent work on the evolution of institutions as the individually unintended collective result of individual actions that are rationally optimizing their own target functions in the presence of limited information includes Refs. 18, 23 and 16. For a pertinent case study, see for instance Ref. 6. This approach has been cast into the framework of game theory by P. Young [30]. Young's setting makes the assumptions that players are randomly drawn from large populations, that their probability of interactions depends on exogeneous factors like spatial proximity, and that they try to act rationally on the basis of the limited information that they have available. Whereas the second point is not elaborated upon in our model, the other two constitute also a basis for our contribution. We should point that in our model, there is no room for bargaining. While this restricts its applications to certain economic situations, it is essential for simplifying the strategy space available to the individual players so that their rational strategy choice behavior can be investigated more easily. Nevertheless, our scenario can be extended to include more complex forms of social interaction, like punishment [11].

Also, Young [30] considers the effects of random perturbations caused by unpredictable exogenous events or agent behaviors. Although we do not elaborate upon all these aspects here, this can also be naturally incorporated into our model, and of course it also relates to the discussion of the stability concepts in the formal analysis [3, 29].

4. Adapting to Random Distributions

Let $k_b(n)$ and $k_s(n)$ be the bids of a buyer and a seller that interact at time $n \in \mathbb{N}$. If the seller chooses his bid k with probability $p(k)$, then his expected gain at time n is

$$\sum_{k=0}^{K} k \; p(k) \; p(k \leq k_b(n)), \tag{1}$$

because he gains k if his bid is not larger than that of the buyer he encounters.

If the populations are sufficiently large (compared to the generation length), the chance that our seller encounters the same buyer repeatedly is small and can be neglected. In particular, this implies that the buyer's bids will not reflect previous bids of that seller, but rather constitute a response to the action of other sellers that are randomly sampled from the population. Assuming that our seller is not in contact with his fellow seller, the action of the buyer will then be random for him, i.e. $k_b(n)$ can be treated as a random number. The underlying distribution will in general not be stationary in time because the buyers may also learn from their previous encounters and adapt their bids correspondingly.

(1) leads to the following simple observation. If we consider $k_b(n)$ as a random variable, the seller should find that response $0 \leq k_s \leq K$ which satisfies

$$k_s = \operatorname{argmax} k \; p(k \leq k_b(n)). \tag{2}$$

If that value is unique, it is the best strategy to always play that number, because (1) is a convex combination of the different k values since $\sum_k p(k) = 1$ as we are dealing with probabilities. If that value is not unique, he can choose any convex combination of the maximizers.

Given the probability distribution for $k_b(n)$, we can then determine the optimal bid k_s (of course, the crucial point is that this probability distribution for the buyer's bids is not known to the seller, and so, in general, he will not be able to identify that optimal value, but nevertheless it is instructive to compare possible strategies with the optimal one). If the distribution for $k_b(n)$ is a Dirac distribution with the value k_0, i.e. if the buyer always selects the same value, k_0, then obviously the seller's best choice is to utilize the same value. If the values of $k_b(n)$ are distributed uniformly, i.e. $p(k_b(n)) = \frac{1}{K+1}$ for $k_b(n) = 0, \ldots, K$, then the expected gain for the seller's choice k is $\frac{k(K+1-k)}{K+1}$, where the optimal value is $k = \frac{K+1}{2}$, which leads to the expected gain $\frac{K+1}{4}$. This then is better than any other response. For example, if the seller also chooses his bid k randomly with the same uniform distribution, his expected gain is

$$\frac{1}{K+1} \sum_{k=0}^{K} k \; p(k \leq k_b(n)) = \frac{1}{K+1} \sum k \frac{K+1-k}{K+1} = \frac{K(K+2)}{6(K+1)}, \tag{3}$$

which is smaller than $\frac{K+1}{4}$. This also applies if the seller chooses as his bid at time n his opponent's bid $k_b(n-1)$ from the previous round. The reason is that we are assuming here that the buyer's bid is random, and so, when following the buyer's previous bid, the seller simply takes a random value for k. If, in contrast, the seller averages the buyers' previous bids, i.e. chooses

$$k_s(n) = \frac{1}{n-1} \sum_{i=1}^{n-1} k_b(i) \tag{4}$$

[where $k_b(i)$ is the value of his opponent encountered at time i], then, when n gets large, by the central limit theorem, $k_s(n)$ approaches the mean value of $k_b(n)$, assuming that the underlying distribution is stationary. When that distribution is uniform, we end up with the optimal value $k = \frac{K+1}{2}$. So, at least in this simple case, we can deduce that averaging over previous bids is a better strategy than simply taking the value encountered in the last step. [This argument generalizes to certain (but not all) more general distributions for the buyer's bids.]

We also see the following. If seller and buyer start with the same random distribution for their bids and adopt the strategy of copying the previously encountered

opponent's bid, then they will stay with that same random distribution forever. Thus, no progress is made in that case. This is somewhat analogous to persistent miscoordination in iterated two-player games; see e.g. Ref. 4. In fact, in our game, copying the previous opponent's bid is the rule that follows from Cournot adjustment, i.e. from always choosing the best response to the previous round [5]. If, in contrast, both players average, then they are both expected to end up with the same value, $k = \frac{K+1}{2}$, if the time n is sufficiently long. Also, if they both try to optimize according to (2), even though they do not know the other's strategy, they are expected to end up with the same fixed value, which if they both act optimally will be the value $k = \frac{K+1}{2}$, which we had already found to be the optimal against a uniform random strategy. But this already leads us to the issue of learning.

5. Learning

In our setting here, learning consists in using the experience from previous encounters to determine the bid for the present round. The problem then is to utilize the information from those previous encounters in the most efficient manner, without wasting too much effort with useless or disadvantageous trials. Since both populations are adapting, we have a more subtle situation than simply trying to learn an unknown, but fixed probability distribution as for example in statistical learning theory [25].

In principle, the learning strategy should compute the actual bid as a function of all previously encountered opponent bids. However, to learn such a function from experience takes a long time. In fact, that time will be too long if the other population settles more quickly at some value that is advantageous for it. Therefore, it makes sense to do some preprocessing of the experience before trying to figure out the response. A natural such complexity reduction consists in taking some suitable average of the encountered opponent values. We have already seen above that in a simple model situation, this is an asymptotically optimal strategy. Of course, if the two populations adapt at the same time, but perhaps according to different strategies, that analysis is no longer strictly valid, and in certain cases it might be better to use some weighted average, with higher weights for the more recently encountered values. In fact, if both populations employ such an averaging strategy, then giving higher weights to the more recently encountered values can lead to a faster convergence to the optimal value.

Although, as already mentioned, the problem here is more difficult than the one addressed by statistical learning theory, it is nevertheless instructive to consider how the latter would go about it. Here, a seller would try to model the probability distribution utilized by the buyers. The models are taken from some model class Λ parametrized by a parameter α. Let $q(k \leq k_b; \alpha)$ be the probability that for the model corresponding to α, the value k is not larger than the buyer's bid k_b. The underlying assumption on which the seller models the buyer here is that the latter's probability distribution is stationary, i.e. does not depend on n. If the seller then

selects his bid k according to a probability distribution $q(k; \alpha)$, his empirical risk after encountering the buyer's bids $k_b(1), \ldots, k_b(n)$ is

$$R_{\text{emp}}(\alpha) = K - \frac{1}{n} \sum_{i=1}^{n} \sum_{k} k \ q(k; \alpha) \ q(k \leq k_b(i); \alpha). \tag{5}$$

Here, he assumes that those bids $k_b(1), \ldots, k_b(n)$ represent an i.i.d. sample of the buyer's distribution. Of course, in general, this assumption is not valid because the buyers also adapt, but we nevertheless proceed. The seller then chooses that parameter $\alpha(n) \in \Lambda$ for which the empirical risk $R_{\text{emp}}(\alpha(n))$ is minimized. By the same convexity argument as before, instead of selecting k according to some distribution $q(k; \alpha)$, he finds that it suffices to take a single value, $k(\alpha(n))$, i.e. to choose a single-valued distribution. His empirical risk then becomes

$$R_{\text{emp}}(\alpha(n)) = K - \frac{1}{n} \sum_{i=1}^{n} k(\alpha(n)) \ q(k(\alpha(n)) \leq k_b(i); \alpha(n)). \tag{6}$$

In particular, he will then play that value $k(\alpha(n))$ at the next step. In principle, as a heuristic strategy, this should also be useful in the case where his opponents are also adapting, even though the values $k_b(i)$ then no longer represent an i.i.d. sample.

6. Dynamics at the Population Level

As explained, we are less interested in the competition between individual agents within a population than in the relative performance of populations with different types of agents. This relative performance is then gauged by the value of the equilibrium eventually reached via the repeated interactions of the agents from the opposite populations. We can make the following simple observations: If the players of one population always play the value k_0, then the other population has no choice but to adapt to that value. This implies that a population whose members converge faster than the ones from the other camp to a value that is favorable to them will be at an advantage. According to the above analysis, an averaging strategy dampens fluctuations and thereby improves the convergence rate. This speed of convergence, however, does not only depend on the own strategies, i.e. need not reflect an absolute superiority of those strategies, but also reflects their reactions and adaptations to the actions of the members of the other population. Thus, it also depends on the state of the dynamics inside the other population. Let us consider an extreme case: the opposite population plays rather randomly — say, chooses offers from the uniform distribution of integers between 0 and K. The population under consideration, however, is subject to quick and strong selection between its agents. That means (1) that the generation length between the evolutionary steps is very short, perhaps even $= 1$. Thus, each agent has only one encounter, after which his performance is evaluated. That also means (2) that only the very best agents, perhaps only a single best one reproduces, and his identical or slightly mutated offspring constitutes

the next generation. In that case, chances are high (and easily computed) that in each generation there is some agent in each generation that strikes a successful deal with a very advantageous offer; for example, a buyer offering very little may by chance encounter a seller asking even less. Thus, because of the speed of evolution assumed, our population will quickly settle around a very favorable value. If the other population then adapts at all, it is forced to accept that value.

Thus, the speed of evolution in one population and the incoherent state of the other population together lead to an equilibrium value that is very favorable to the first population.

In contrast, for any averaging strategy, the speed of convergence will be slower when playing against a more random opposite camp.

For more complex strategies, a more erratic state of the opposite population may also slow down the own convergence. If a particular response needs to be created for each previous opponent bid (as in the complex one-round opponent strategy described below), then an erratic opponent population forces the players to test many different options, and a player that has already created good responses for many bids may still acquire a low fitness it he is by chance exposed to bid values not yet encountered by himself or his ancestors. Conversely, if the opposite population is rather homogeneous and constant, the players may evolve quickly to seemingly stable states, but this may hide the fact that they do not possess adequate responses to situations that, while possible by the rules of the game, they and their ancestors never experienced.

7. Different Strategies and Simulations

Before presenting the simulation results, we need to introduce some notations for the system parameters and the strategies that will be used in this paper.

First, we present the parameters for the evolutionary scheme of replacing a population of players by a new one composed of possibly mutated members of the present one with a fitness-based selection: (1) generation length (time) — the number of rounds played (time steps) between two consecutive selections (if applicable); (2) selection percentage — the percentage of the players who will be chosen as parents to generate the offspring during the evolutionary process; (3) mutation rate — the rate of random mutation during the evolutionary process.

Next, we list the main strategies investigated, classified on the basis of the types of information they use:

(1) Single-number: players use no information at all; each player chooses a fixed random offer that will be updated through the selection.
(2) Average-previous-opponent: the average of one's opponents' bids in the previous — say, m (limited and usually much smaller than the generation length) — rounds.
(3) For $m = 1$, that strategy is called one-round opponent: each player utilizes the offer of his opponent in the most recent round.

(4) n-round opponent: each player can use the offers his opponents make in the last n rounds (and not only their average) — here, we consider only the values $n = 1$ (which is the previous strategy) and 2, as otherwise the scheme gets computationally too complex and performs too poorly.
(5) Average-all-previous-opponent: the average of one's opponents' bids in all previous rounds (thus, there is no fixed m here).
(6) Average-friend-opponent: the average of one's friends' opponents' bids in the most recent round (here, each player has a certain number of friends within his own population).
(7) Average-all-friend: the average of one's friends' bids in the most recent round (thus, here, in contrast to the previous strategies, no information about the other population is used).
(8) Friendship network (average-successful-friend): the average of one's friends' successful bids in the most recent round (here, information from the other population is used indirectly, but selectively, because their offers decide which of the friends are successful).

For the n-round opponent, for $n > 1$, a scheme is needed to convert the n numbers remembered into a single response. Of course, we could simply take their average, which would then reduce this to strategy 2, but we could also employ some other scheme. One possibility is to evolve a look-up table that lists the responses for any pair of numbers between 0 and K. Similarly, we can also utilize look-up tables for the other strategies except strategy 1, i.e. instead of simply playing the corresponding average, each agent could have a look-up table that specifies a specific response to each number between 0 and K (formally, strategy 1 is also a special case of this, the output of the look-up table simply being reduced to a single number that is uniformly applied to any input). Thus, a strategy comes in two variants — a direct one and another with a look-up table. We call these two variants "simple" and "complex", respectively. As noted, the two-round opponent strategy does not possess a simple version.

First of all, we would like to have a stable setting for our model. The stability of the model may primarily depend on the values of the system parameters, i.e. generation length, selection percentage and mutation rate. The effect of the generation length is somewhat dependent on the complexity of the strategy under consideration. We will go into this issue further later on. Generally, if the selection process takes place, accompanied by evolving the look-up tables, then an appropriate generation length will be any number ranging from 100 to 1000 time steps (rounds of the game). As one knows from evolutionary optimization methods, an appropriate choice of the selection percentage is essential. If it is too small, then the selection will be severe and some potentially good strategies can get eliminated too easily. If it is too large, then the selection will be very loose and the optimization will take much longer. According to our numerous simulations, an effective selection

percentage will be 0.5. For the random mutation rate, a good choice will be 1–5%, and in our simulations it is set to be 1%. Our population size is always 400.

7.1. *The effects of the generation length*

The generation length, i.e. the number of rounds played before an evolutionary update, expresses the relation between the time allocated to learning in the more complex strategies and the evolutionary adaptation. A short generation length means that individual agents have little time to improve their performance on the basis of memory and learning, but rather are evaluated according to their short time performance. In other words, their experience is quickly transferred to the next generation. That generation can then explore new responses not on the basis of systematic learning, but on the basis of random mutations.

We can then simply check this issue in our simulations by letting two populations with the same strategy space, but with different generation lengths, play against each other, and see which one performs better.

For the single-number strategy, it then turns out that the minimal generation length, 1, is optimal. This is rather obvious because a more quickly evolving population should have an advantage. This is confirmed by the simulation results. For more complex strategies such as the one- or two-round opponent strategy or the friendship network strategy, the situation is not so clear, as there is a tradeoff between individual improvements based on more experience and the speed of the evolutionary search. The simulations results are not yet conclusive. In some simulations, there is a small chance for the two-round opponent strategy to gain a slight advantage over the one-round opponent strategy if the generation length is as long as 20,000 steps or even longer. Note that for shorter generation lengths, we have demonstrated in Ref. 10 that the one-round strategy is superior to the two-round because the latter takes too long to evolve. Further simulations with very long generation times are required.

7.2. *Efficient information use*

A simple analysis uses the Shannon entropy [21],

$$S = -\sum_{k=0}^{K} p(k) \log_2 p(k), \tag{7}$$

where $p(k)$ is the probability (frequency) that offer k has been chosen. In fact, this will generally depend on the time step n, and so we should rather write $p(k,n)$ [which may approach a stationary $p(k)$ as $n \to \infty$].

S is maximal for a uniform distribution of the values of k and becomes 0 when only a single value of k is played. In other words, the evolution of S expresses how quickly a population reaches a unique response. One may argue that (7) expresses uncertainty and that therefore a fast decrease of this entropy corresponds to fast

utilization of information. Figure 1 gives some simulation results. In particular, populations with a fast decrease of this entropy are more successful.

The simulations behind Fig. 2 show that the complex average-all-previous-opponent strategy performs better than the complex one-round opponent strategy. One could argue that this should be so because the former utilizes more information in each step than the latter. A simpler reason is that averaging reduces fluctuations and therefore speeds up convergence. Further figures (Figs. 4 and 5) demonstrate that simple strategies converge faster and perform better than complex ones. However, the simple one-round opponent strategy performs poorly when compared with strategies utilizing some nontrivial averaging. This is in line with our mathematical analysis above.

7.3. *Ranking different strategies*

So far, we have investigated many different types of strategies. It is of interest to check whether we can consistently rank their performances. The basis for such a ranking is of course a pairwise comparison, i.e. we let one population — say, the buyers — employ one type of strategy and the sellers another. Here, we need to fix the other parameters, like generation length or mutation rate, to have a *ceteris paribus* comparison, even though we realize that in principle the ranking could change with different parameter values. In the simulations presented below, the generation length is 1000, the selection percentage 0.5 and the random mutation rate 1%.

In Ref. 10, some of the complex strategies have been compared. The one-round opponent strategy can beat the two-round opponent strategy because the former settles down to equilibrium more quickly. The performance of the friendship network strategy is comparable to that of the one-round opponent strategy, neither of them showing consistent superiority. The friendship network strategies do not prefer any specific type of network topology (keeping the average degree of the networks fixed).

For a more systematic investigation, however, we should compare different types of averaging strategies, in particular the average-previous-opponent and average-friend-opponent strategies. In the former, the average is performed over the player's own opponents' offers in the previous — say, five — rounds, whereas in the latter he averages his friends' opponents' bids in the last round. In the average-friend-opponent strategy, one can include the player's own experience, i.e. his own opponent's bid in the last round. Because the opponents are randomly taken from the opposite population, for sampling purposes this makes no difference, as long as the average is computed from the same number — say, five — of bids from the opposite camp. As the network topology might affect the speed of convergence of the averaging scheme, this might have some effect here, however. In any case, for a valid comparison between strategies, we assume that the numbers of offers taken for both averaging procedures are the same — say, five.
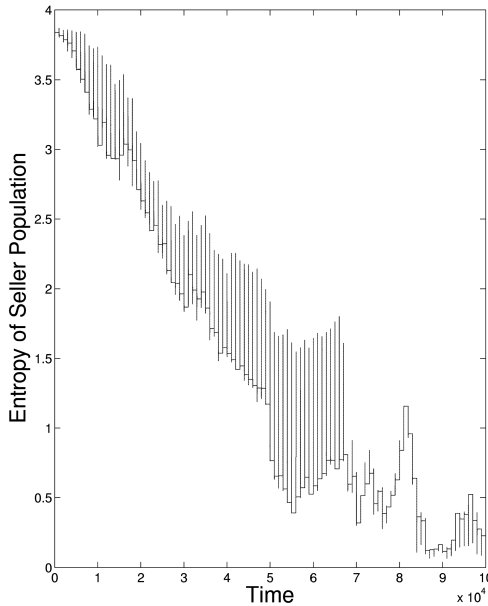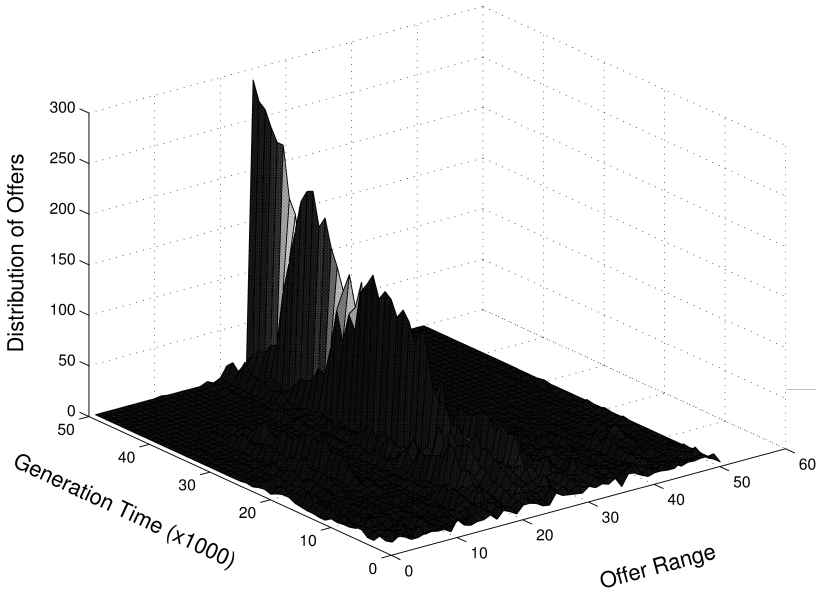
Fig. 1. The top panel shows the time evolution of distribution of offers made by the sellers during the first 50 generations (1000 rounds per generation). Here both populations use the one-round opponent strategy. As shown, the peak where the most favorable offer appears becomes higher as the evolution continues. The final value of the peak in this figure is 300, which shows the number of players who have bid 25. The bottom panel shows the time evolution of the entropy related to the distribution of sellers' offers.
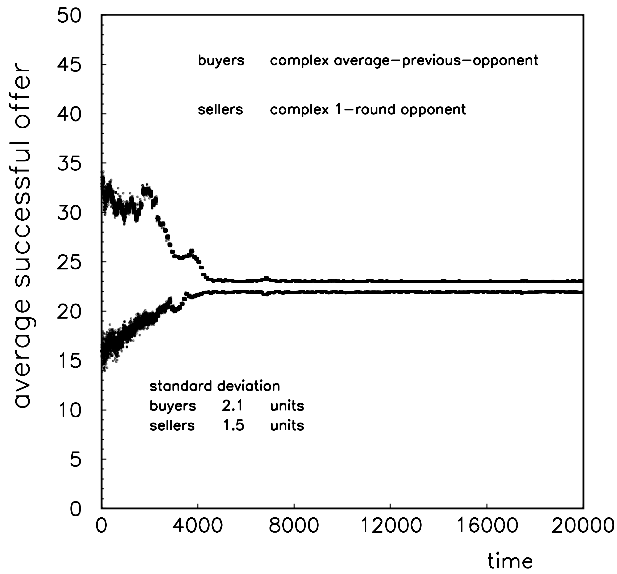
(a)
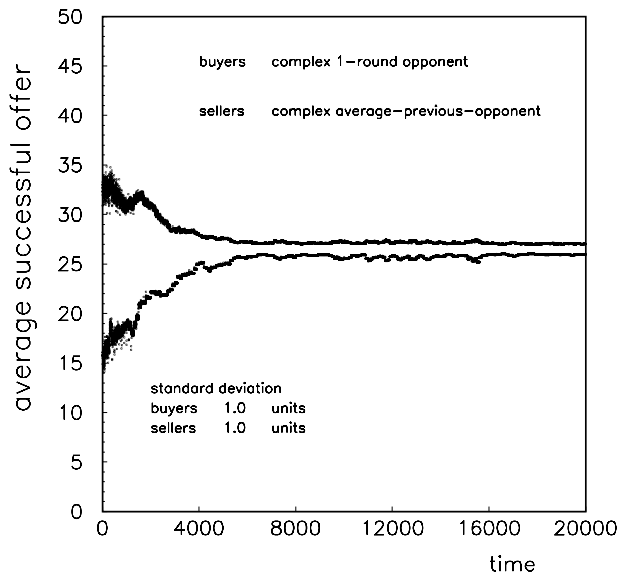
(b)

Fig. 2.   The complex average-all-previous-opponent strategy, where the player bids according to all his previous encounters' offers by taking an average, versus the complex one-round opponent strategy, where the player recalls only his most recent interaction ("complex" here means developing look-up tables). In both (a) and (b), the population with the average-all-previous-opponent strategy is doing better. Learning is quick and the equilibrium is rather stable. The standard deviation is calculated from 100 samples with different realizations.
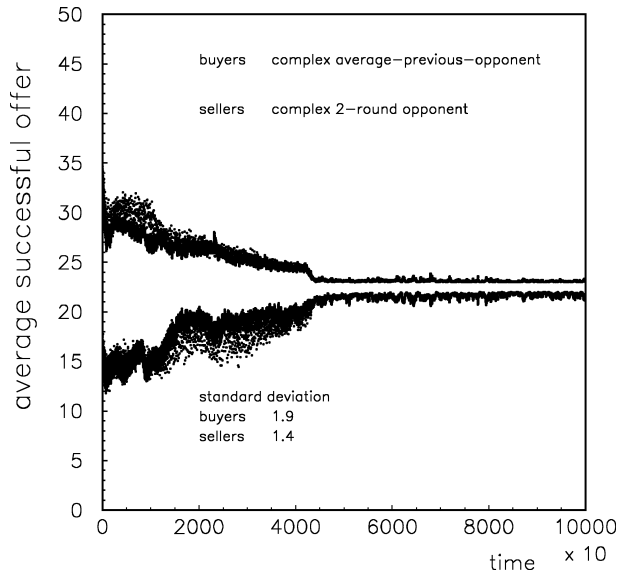
(a)
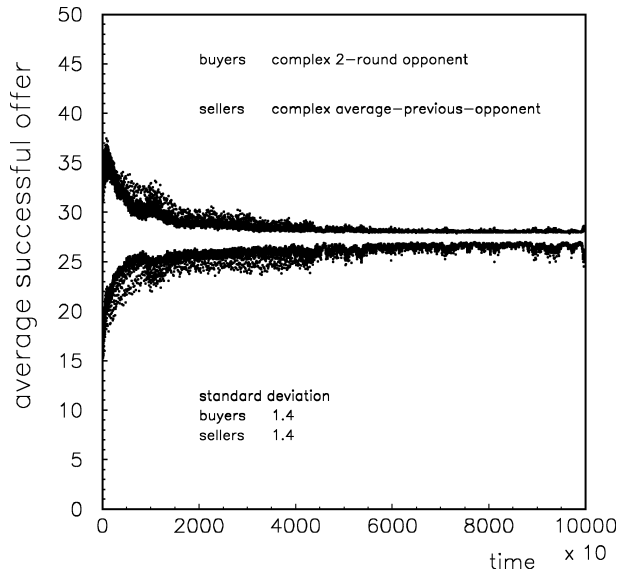
(b)

Fig. 3. The complex average-all-previous-opponent strategy versus the two-round opponent strategy. In both (a) and (b), the averaging strategy is superior. Compared to the situation in Fig. 2, the optimization takes much longer and there also exist some fluctuations even after the equilibrium has been reached. This occurs mainly due to the excessive information embedded in the two-round opponent strategy.
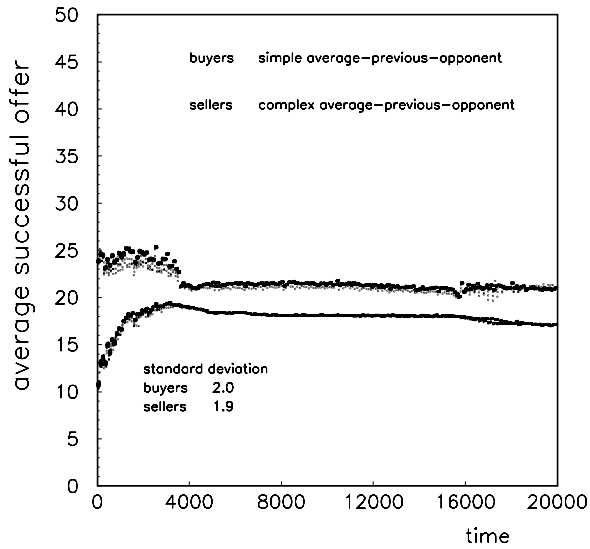
(a)
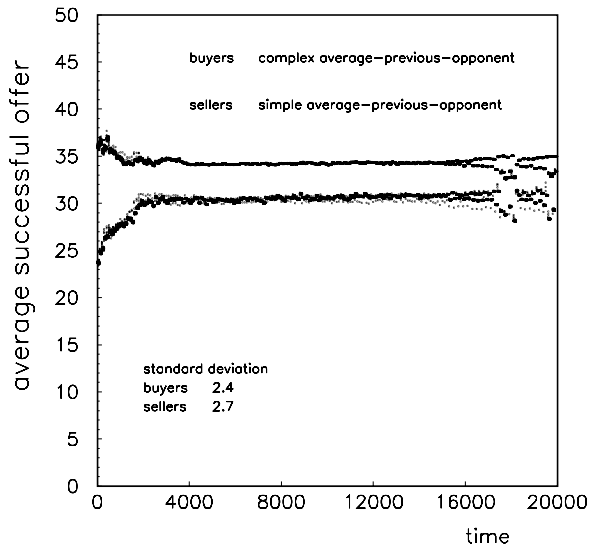
(b)

Fig. 4.   The simple versus the complex average-all-previous-opponent strategy. The simple averaging strategy is superior to the complex one. The median fee, namely the gap, is still somewhat high in both (a) and (b). This happens mainly because the players who employ the simple averaging strategy can converge more quickly and set an early advantage. The players who use the complex averaging strategy have to adapt accordingly and are forced into a disadvantageous situation. Learning is quick because of the convergence due to the simple averaging.
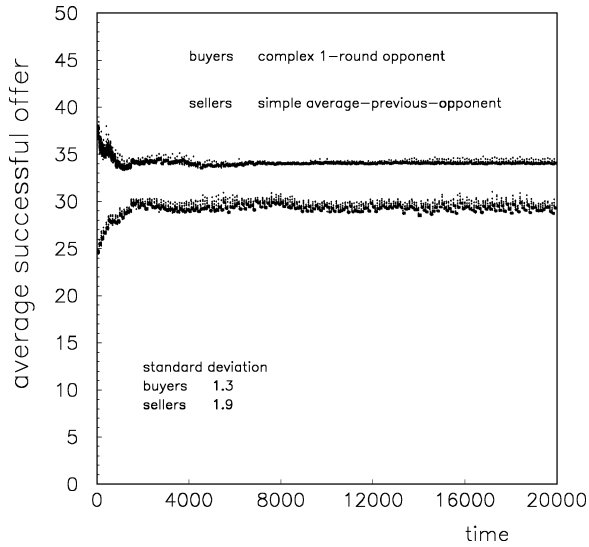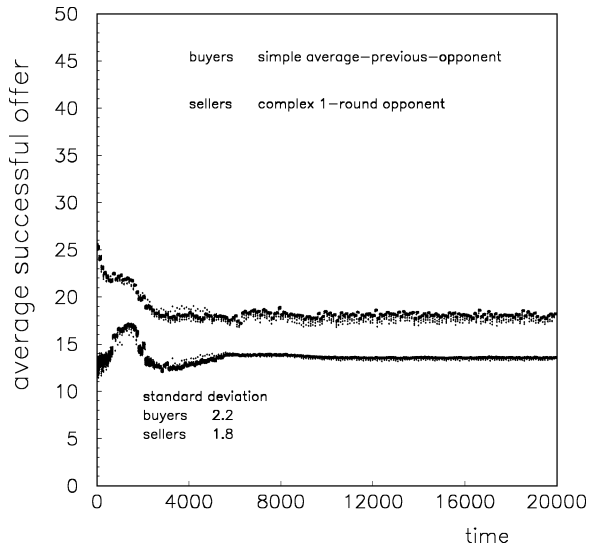
Fig. 5.   Comparison between the simple average-all-previous-opponent strategy and the complex one-round opponent strategy. The former is better.

We first compare the simple average-previous-opponent to the simple average-friend-opponent, i.e. the strategies not employing look-up tables, but taking the computed average directly as the next own bid. The first one uses a somewhat larger sample, because it takes the bid of one member of the opposite population at five different times, i.e. it takes a spatiotemporal average; whereas the second
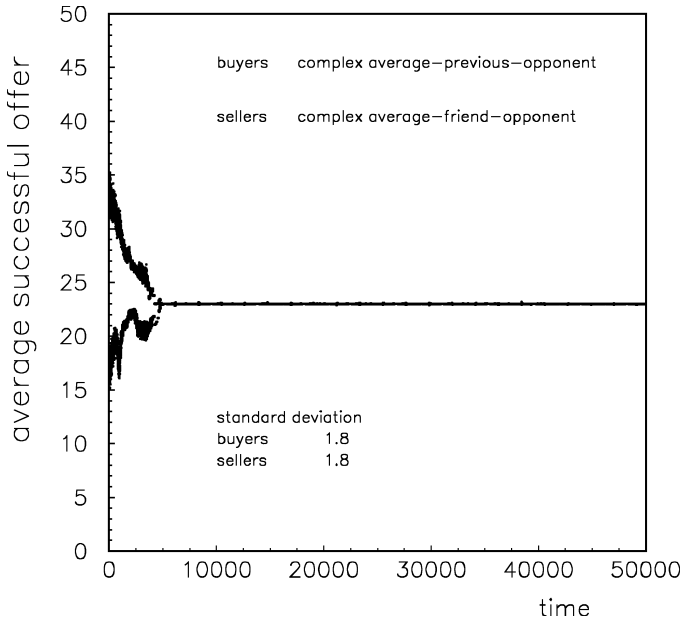
Fig. 6.   The complex average-previous-opponent strategy versus the complex average-friend-opponent strategy. The former can do slightly better than the latter, mainly because it has a slight larger sample by doing spatiotemporal averaging and the latter is doing only spatial averaging.

averages over five bids taken at the same time, i.e. it takes only a spatial average. In any case, both averaging strategies quickly converge to their equilibrium, the population average of the opposite population, typically 25. It also appears that the average-previous-opponent converges slightly more slowly than the average-friend-opponent, because the latter uses a more recent sample from the opposite camp.

When we look at the complex strategies, i.e. where look-up tables are used to convert the computed average into a response,we see from Fig. 6 that in general the average-previous-opponent can be slightly better than the average-friend-opponent. Thus, the larger sample space, even though it uses partly outdated information, can yield an advantage.

We now turn to the average-successful-friend strategy. Here, each player takes the average only of those offers of his friends and himself from the last round that have led to a successful deal. (If none of those offers is successful, then a random offer between 0 and K is chosen.) (This average-successful-friend strategy is the friendship network strategy examined in Ref. 10.) At the beginning of the game, some randomness is introduced due to the low success rate. Also, apparently, choosing only successful offers for averaging makes the players more and more timid in making their offers. This is the reason why the average-previous-opponent strategy
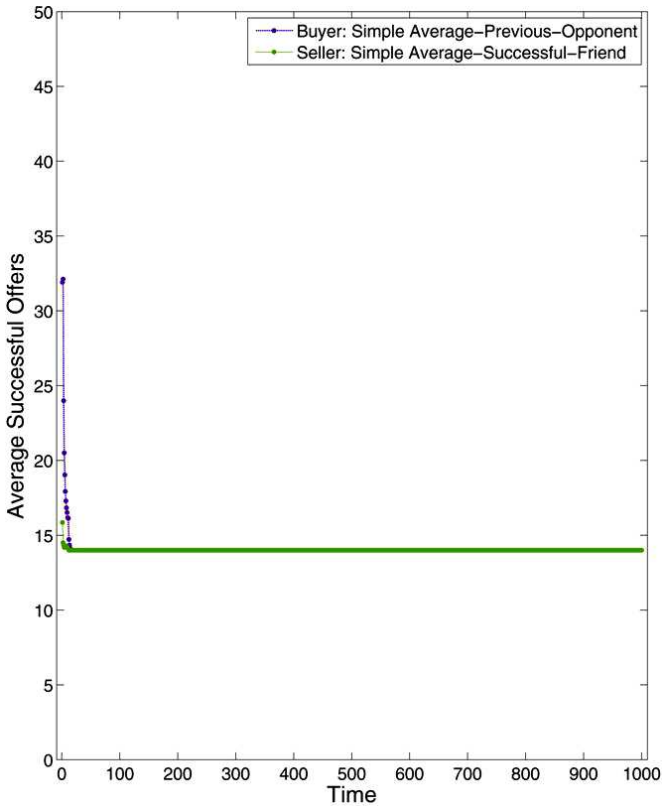
Fig. 7. The simple average-previous-opponent strategy versus the simple average-successful-friend strategy. The former is superior. Distinguishing the friends by success makes the players too cautious. The standard deviation of equilibrium values for both buyers and sellers is zero.

is superior to the average-successful-friend strategy, in both the simple and the complex setting. In Fig. 7, the buyers who use the simple average-previous-opponent strategy can achieve the equilibrium value 14 by adapting to the sellers who use the average-successful-friend strategy. (A coarse estimate would in fact expect the even lower equilibrium value 12.5.) Figure 8 presents the competition between the two strategies with look-up tables. Now the difference in the performance between the two strategies is not that large, compared to the case without look-up tables. This is partly due to the fact that the players are learning. We also notice that it takes the average-previous-opponent a little longer than the average-successful-friend to reach the equilibrium.

In the average-all-friend strategy, a player does not distinguish whether his friends are successful or not, but simply averages all their bids. The simple average-all-friend strategy matches the simple average-previous-opponent strategy, with both converging to 25. This happens because the two now make use of the same trivial random distribution. Not surprisingly, the complex average-previous-opponent
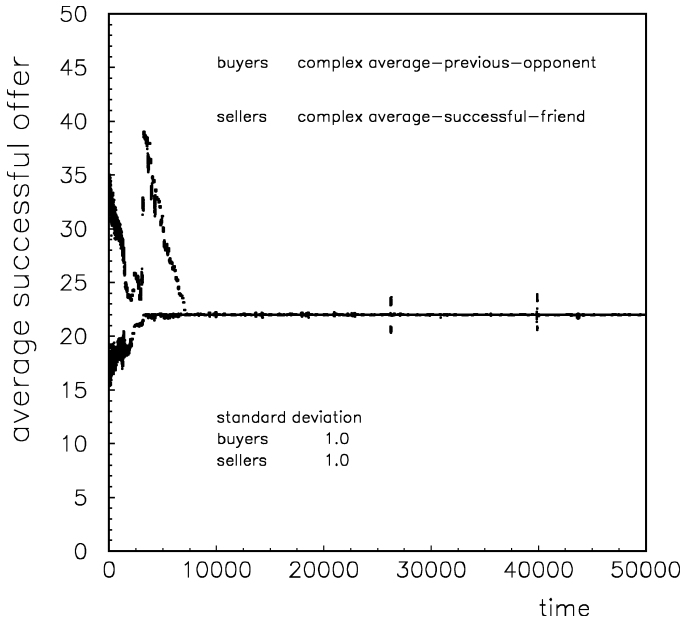
Fig. 8.   The complex average-previous-opponent strategy versus the complex average-successful-friend strategy, with the former defeating the latter. Here the advantage is not significant as in Fig. 7, primarily due to effective learning.

strategy performs better than the complex average-all-friend strategy; see Fig. 9. The reason, as above, is that the former is using a larger sample space.

We have also compared the efficiency of different spatial averaging strategies, namely average-friend-opponent, average-successful-friend and average-all-friend. The comparison between the last two is rather straightforward, with average-all-friend prevailing over average-successful-friend. The average-friend-opponent strategy was found to be better than the average-successful-friend strategy, with and without look-up tables. Figure 10 shows an example of our simulations. It is interesting to see that the average-friend-opponent strategy is performing nearly equivalently to the average-all-friend strategy, without and with look-up tables (Fig. 11). This observation again confirms that using less selective information can be advantageous. This is not surprising, since we have already observed that the one-round opponent strategy can beat the two-round opponent strategy.

We can now rank the different strategies. The best strategy should be simple average-all-previous-opponent. We have found that average-previous-opponent is nearly as good as average-all-previous-opponent if the number of rounds for averaging is not too small. Hence the second rank consists of complex average-all-previous-opponent and average-previous-opponent. The third position is taken by average-all-friend and average-friend-opponent. At the fourth position, we would put one-round opponent and average-successful-friend. The lowest position belongs to two-round opponent. There is yet one strategy that needs to be placed: the
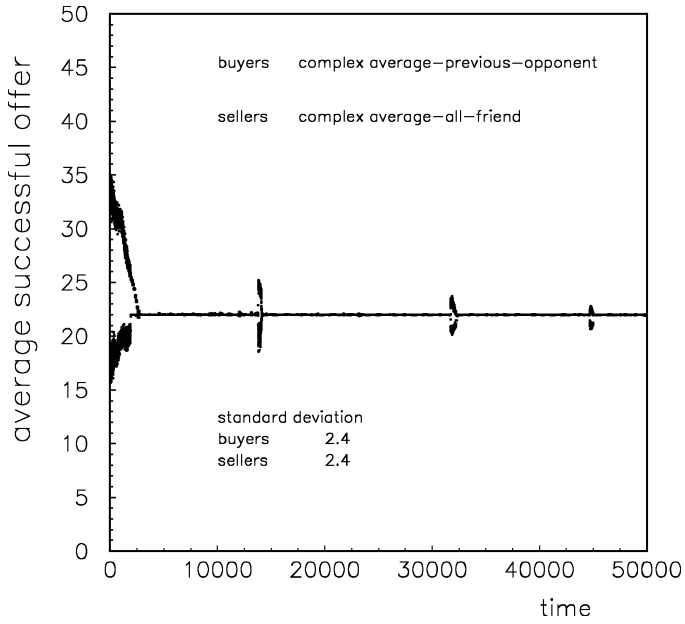
Fig. 9.   The complex average-previous-opponent strategy versus the complex average-all-friend strategy, with the former being slightly better.
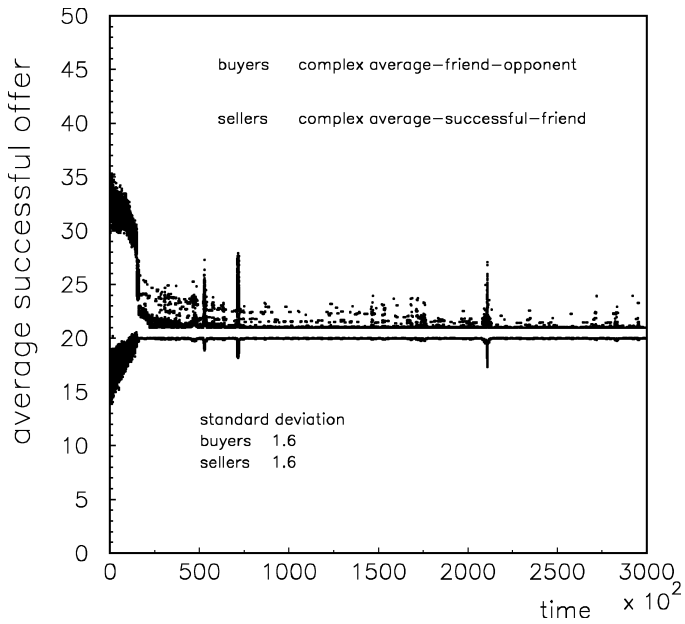

Fig. 10.   The comparison between the complex average-friend-opponent strategy and the complex average-successful-friend strategy. The former is better, which indicates again that selecting only successful friends makes the players too timid.
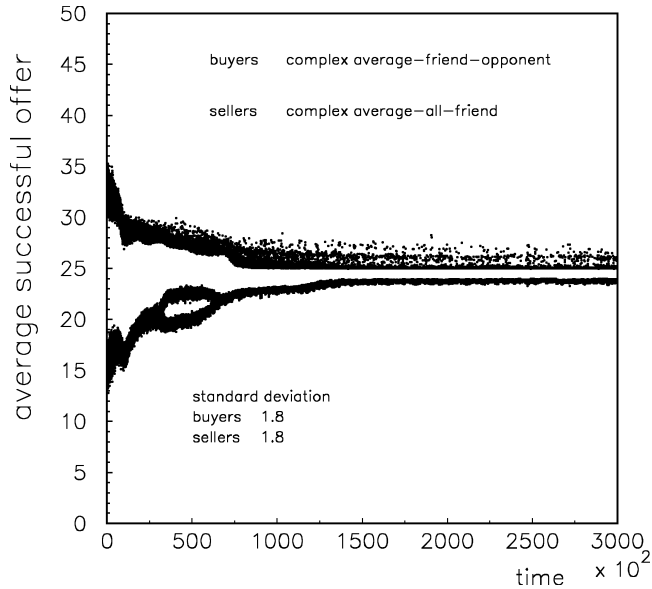
Fig. 11.   The complex average-friend-opponent strategy is almost equivalent to the complex average-all-friend strategy. Compared to what has been shown in Fig. 10, using less selective information is not inevitably disadvantageous.

single-number strategy. This last strategy does not use any information but can still be favored in the competition with other, more complex ones that use more information.

From the above ranking we see that whether a strategy is good or bad is not completely decided by how much information it has used. Rather, a good strategy should be capable of processing the information more efficiently and thus setting an early advantage as quickly as it can. In our simulations, we find that a weighted averaging strategy is not doing better than the normal averaging strategy. The weighted averaging is perhaps more reasonable by assigning the most probable offers more chances but its adaptation also takes longer. Being simpler and more flexible is also good for a strategy. In this regard, the simple averaging strategy that does not need to evolve look-up tables can beat the complex one that does; not to mention the fast that the single-number strategy can beat a lot of more complex strategies.

8. Conclusion

We have investigated an iterated game played between members of opposite populations. The individual optimization leads to a population dynamics that determines the final equilibrium reached. Equipping the members of the two populations with different strategic options will in general lead to an equilibrium that is more favorable to one of the two populations. A set of strategies that is good for a population

when consistently employed by all its members incorporates quick and efficient use of the available information without a long learning phase, i.e. rather than foregoing careful optimization when it takes too long to converge. Also, a population with bolder players does better than one with more timid players.

It remains for us to investigate the dynamics within populations more systematically when the individual players have different strategic options. The ones that would cause advantageous effects at the population level might themselves be disadvantaged inside their population and therefore get eliminated by the evolutionary scheme, which would then also cause a disadvantage for the population.

Acknowledgments

References

[1] Amann, E. and Hofbauer, J., Permanence in Lotka-Volterra and replicator equations, in: Ebeling, W. and Peschel, M. (eds.), Lotka-Volterra approach to cooperation and competition in dynamic systems (Akademie-Verlag, 1985).

[2] Cressman, R., The stability concept of evolutionary game theory (Springer, 1992).

[3] Foster, D. and Young, H. P., Stochastic evolutionary game dynamics, *Theor. Pop. Biol.* **38**, 219–232, 1990; correction in *Theor. Pop. Biol.* **51** (1997) 77–78.

[4] Fudenberg, D. and Kreps, D., Learning mixed equilibria, *Games Econ. Beh.* **5** (1993) 320–367.

[5] Fudenberg, D. and Levine, D., The theory of learning in games (MIT Press, 1998).

[6] Greif, A., Contract enforcability and economic institutions in early trade: The Maghribi trader's coalition, *Amer. Econ. Rev.* **83** (1993) 525–548.

[7] Hofbauer, J., Schuster, P. and Sigmund, K., A note on evolutionarily stable strategies and game dynamics, *J. Theor. Biol.* **81** (1979) 609–612.

[8] Hofbauer, J. and Sigmund, K., *The Theory of Evolution and Dynamical Systems* (Cambridge University Press, 1988).

[9] Hofbauer, J. and Sigmund, K., *Evolutionary Games and Population Dynamics* (Cambridge University Press, 1998).

[10] Jost J. and Li, W., Individual strategies in complementarity games and population dynamics, *Physica A* **345** (2005) 245–266.

[11] Li, W., Cai, X. and Wang, Q. A., Punishment in a complementarity game, *Physica A* **364** (2006) 403–412.

[12] Mandeville, B., The fable of the bees: or, private vices, publick benefits. With a commentary critical, historical, and explanatory, (ed.) F. B. Kaye, Vol. 2 (Oxford, 1924, 1957).

[13] Maynard, J., Smith and Price, G., The logic of animal conflict, *Nature* **246** (1973) 15–18.

[14] Menger, C., Grundsätze der Volkswirtschaftslehre (Wien, 1871).

[15] Nash, J., Equilibrium points in n-person games, *Proc. Nat. Acad. Sci.* **36** (1950) 48–49.

[16] North, D., *Institutions, Institutional Change, and Economic Performance* (Cambridge University Press, 1990).

[17] Samuelson, P., *Economics* (McGraw Hill, 1973).
[18] Schotter, A., *The Economic Theory of Social Institutions* (Cambridge University Press, 1981).
[19] Schumpeter, J., *The Theory of Economic Development: An Inquiry into Profits, Capital, Credits, Interests and the Business Cycle* (Oxford University Press, 1934).
[20] Selten, R., A note on evolutionarily stable strategies in asymmetric animal conflicts, *J. Theo. Bio.* **84** (1980) 93–101.
[21] Shannon, C. E. and Weaver, W., *The Mathematical Theory of Communication* (University Ill. Press, 1949).
[22] Smith, A., *An Inquiry into the Nature and the Causes of the Wealth of Nations* (London, 1789).
[23] Sugden, R., The evolution of rights, cooperation, and welfare (Basil Blackwell, 1986).
[24] Taylor, P., Evolutionarily stable strategies with two types of player, *J. Appl. Prob.* **16** (1979) 76–83.
[25] Vapnik, V. N., *Statistical Learning Theory* (Wiley, 1998).
[26] Vega-Redondo, F., *Economics and the Theory of Games* (Cambridge University Press, 2003).
[27] von Neumann J. and Morgenstern, O., *Theory of Games and Economic Behavior* (Princeton University Press, 1944).
[28] Weibull, J., *Evolutionary Game Theory* (MIT Press, 1995).
[29] Young, H. P., The evolution of conventions, *Econometrica* **61** (1993) 57–84.
[30] Young, H. P., *Individual Strategy and Social Structure* (Princeton Universtiy Press, 1998).