Equilibrium Selection in Alternating-Offers Bargaining Models – The Evolutionary Computing Approach

D.D.B. van Bragt E.H. Gerding J.A. La Poutré
{bragt,egerding,hlp}@cwi.nl

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

ABSTRACT
A systematic validation of evolutionary techniques in the field of bargaining is presented. For this purpose, the dynamic and equilibrium-selecting behavior of a multi-agent system consisting of adaptive bargaining agents is investigated. The agents' bargaining strategies are updated by an evolutionary algorithm (EA), an innovative computational method to simulate collective learning in societies of boundedly-rational agents. We show that game theory can be used successfully to interpret the equilibrium-selecting behavior observed in computational experiments with adaptive bargaining agents. Deviations from classical game theory, for instance the occurrence of relaxation oscillations in the ultimatum game, are however observed in several experiments.

JEL: C61, C63, C72, C78, D83
Keywords and Phrases: evolutionary algorithms, evolutionary game theory, agent-based computational economics, alternating-offers bargaining, bounded rationality
Note: This paper has been presented at the Sixth Computing in Economics and Finance Conference (CEF'2000) (Barcelona, Catalonia, Spain, July 6-8, 2000) and the First World Congress of the Game Theory Society (GAMES'2000) (Bilbao, Spain, July 24-28, 2000). This research has been performed within the framework of the project "Autonomous Systems of Trade Agents in E-Commerce", which is funded by the Telematics Institute in the Netherlands.

1. Introduction

Recently, interest in the development of trading and negotiating agents has surged among economists and computer scientists [4]. A nice example of the potential of automated negotiation is given in [5]. This paper describes a system in which a utility agent (acting on behalf of an electricity company) is negotiating with consumer agents to prevent excessive peaks in the demand for electricity. Another example is the agent-based heating system of the Xerox company. In this climate control system each agent controls an office thermostat and the allocation of resources is market-based. Practical applications of distributed multi-agent systems are surveyed in [24, Ch. 9].

The rapid establishment of a global communication network (in the form of the Internet) together with the development of standard negotiation protocols [15] will certainly result in a fast proliferation of systems of this kind. The complexity of large multi-agent systems increases strongly, however, if the negotiating agents are not using fixed decision rules but adapt their strategies to deal with changing opponent strategies and changing user preferences. Two important and fundamental questions should therefore be raised: (i) which complex dynamic behavior will emerge in this kind of complex adaptive systems, and (ii) to which state will these systems converge over time (if a stable steady state is reached at all).

We address these two issues in a computational setting by using evolutionary algorithms (EAs). EAs, inspired by Darwin's theory of evolution, are an attractive tool to model collective learning in societies of boundedly-rational agents [14, 6]. In an evolutionary setting, the adaptive agents learn in three different ways: (i) learning by imitation (by reproduction and selection of successful strategies),

(ii) communication and exchange of strategic information (by recombining or "crossing over" genetic information), and (iii) random experimentation (by "mutating" their strategies).

Oliver [12] was the first to demonstrate that a system of adaptive agents can learn effective negotiation strategies. Computer simulations of both distributive (i.e., single-issue) and integrative (i.e., multiple-issue) "alternating-offers" negotiations are presented in [12]. Binary-coded strings represent the agents' strategies. Two parameters are encoded for each negotiation round: a threshold which determines whether an offer should be accepted or not and a counter offer in case the opponent's offer is rejected. These strategies were then updated in successive generations by a genetic algorithm (GA).

More elaborate strategy representations are proposed and evaluated in [9]. Offers and counter offers are generated in this model by a linear combination of simple bargaining tactics (time-dependent, resource-dependent, or behavior-dependent tactics). As in [12], the parameters of these different negotiation tactics and their relative importance weightings are encoded in a string of numbers. Competitions were then held between two separate populations of agents, which were simultaneously evolved by a GA.

We intend to bridge the gap between computer experiments as performed in [12, 9] and the analysis of bargaining by game theorists [20, 17, 13, 11]. This connection is not far-fetched. Consider first how agents in the computer experiments learn to bargain in an evolutionary model. Initially, agents will typically use a random strategy. As a consequence, many different paths through the game tree will be explored (i.e., many subgames will be sampled). Only the agents with relatively successful strategies in many different subgames will be selected as parents for the next generation of agents. In each successive generation, this process of variation and selection is then repeated and more and more robust strategies evolve in the long run.

Now consider the key equilibrium concept used by game theorists to analyze extensive-form games:[1] the subgame-perfect equilibrium (SPE) [18, 19]. Two strategies are in SPE if they constitute a Nash equilibrium in any subgame which remains after an arbitrary sequence of offers and replies made from the beginning of the game. Rubinstein successfully applied this notion of subgame-perfection to bargaining games [17]. His main theorem states that the infinite-horizon alternating-offers game has a unique SPE in which the agents agree immediately on a deal.[2]

Recent work in evolutionary game theory illustrates the link between subgame-perfectness and (mathematical) evolutionary models. A nice example is given by Binmore et al. [3]. They consider an evolutionary model in which finite-state machines play Rubinstein's infinite-horizon bargaining game against each other. For such a system, Binmore et al. prove that if both agents use a so-called "modified evolutionarily stable strategy" a Nash equilibrium is constituted in which immediate agreement is reached. Furthermore, each agent's share of the surplus is bounded between the shares received by the two agents in the SPE of the infinite-horizon game.

Hence, the results in [3] indicate that agents with a bounded rationality may actually display subgame-perfect behavior in the evolutionary alternating-offers game. Our computational experiments confirm this point. Moreover, we encounter phenomena beyond the reach of classical game theory. For example, if the agents' discount factors are very small (i.e., when time pressure to reach an agreement is extremely large) strongly nonlinear behavior is occasionally observed (depending on the specific evolutionary selection scheme). If discount factors are large, on the other hand, (i.e., when time pressure is weak) the finite horizon of the game is not fully exploited by the agents. Significant deviations from game-theoretic predictions are also observed if the agents discount their payoffs at a different rate.

The remainder of this paper is organized as follows. Section 2 gives an overview of the bargaining models that we investigate. A description of the setup of the computational experiments is then given in Section 3. Sections 4-8 provide an overview and discussion of the main results.

The finite-horizon variant of Rubinstein's alternating-offers model [17] is analyzed in Sections 4

[1] That is, games with a tree structure [2].

[2] The finite-horizon variant of Rubinstein's game (which we use in our computer simulations) has been analyzed earlier by Ståhl [20].

through 6. In Section 4 the agents have symmetrical time preferences (i.e., their discount factors are identical). If the agents' discount factors approach zero in this model (i.e., when the bargaining surplus completely vanishes after one round), this game becomes payoff-equivalent with the well-known ultimatum game. This limiting case is studied in detail in Section 4.1. Another extreme case is obtained if the agents' discount factors approach unity (i.e., when the bargaining surplus remains constant over time). This situation is investigated in Section 4.2. The general case (i.e., discount factors in between zero and unity) is analyzed in Section 4.3. The robustness of the experimental results with respect to changes in the evolutionary model is assessed in detail in Section 5. The influence of asymmetric time preferences (i.e., unequal discount factors) is evaluated in Section 6.

Section 7 considers a variant of Rubinstein's model in which there exists a risk of premature breakdown during the negotiations. Extending this model, the behavior of risk averse agents is examined in Section 8. Section 9 concludes.

2. Alternating-Offers Bargaining Models

We consider several variants of the alternating-offers game in this paper, ranging from the very simple ultimatum game in Section 4.1 to a multiple-stage game with breakdown and risk averse agents in Section 8. Before presenting these models in more detail, we like to mention that the existing literature on alternating-offers bargaining [13, 11] mainly considers infinite-horizon games. However, in our computational experiments bargaining obviously cannot continue for arbitrary lengths of time. We therefore perform a game-theoretic analysis of finite-length games in Appendix 1. For games with a very long horizon, correspondence with existing results for infinite-length games is shown.

It is also important to note that we assume in the game-theoretic analysis of Appendix 1 that the bargaining agents behave fully rational and have complete information (for instance about their opponents' preferences). Both assumptions are obviously not valid for the evolving agents in our computational experiments (who learn by trial-and-error instead of abstract reasoning). However, the (subgame-perfect) equilibrium behavior of fully rational agents will serve as a useful theoretical benchmark to interpret the behavior of the boundedly-rational agents considered in Sections 4-8.

The ultimatum game (Section 4.1) The ultimatum game is the most simple bilateral bargaining game with perfect information. Agents bargain over the partitioning of a given bargaining surplus. Without loss of generality, we can set the size of this surplus equal to unity. We also assume that the agents have opposite preferences. One agent (which is denoted in this paper as "agent 1") then starts the play by demanding a certain fraction x_1 of the surplus. The other agent ("agent 2") has only two options: he can either accept or reject agent 1's proposal. In case the offer is accepted agent 1 receives x_1 and agent 2 the remainder (i.e., $1 - x_1$). Both agents receive nothing if the offer is rejected.

Multiple-stage games (Section 4.2) In the multiple-stage variant of the ultimatum game, the agents again bargain over the partitioning of a constant surplus. Offers are now made at discrete points in time: namely, at times $t = 0, ..., (n-1)$, where n is the maximum number of stages of the bargaining game.[3] The two agents bargain in an alternating fashion. At $t = 0$, agent 1 makes an offer. Agent 2 then accepts or rejects this initial offer. If the initial offer is rejected, agent 2 makes a counter offer in the next round (at $t = 1$). This alternating process of making proposals then continues until an offer is accepted or until the bargaining deadline is reached (at $t = n$). If no agreement has been reached before the deadline (that is, for $t < n$) both agents receive nothing.

Model with time preferences (Sections 4.3 through 6) Agents are now under time pressure because they prefer to reach an agreement early. Following Rubinstein [17] we model the time preferences of agent $i = 1, 2$ with a discount factor δ_i, with $0 < \delta_i < 1$. In case of an agreement, agent i's discounted payoff is equal to $x_i \delta_i^t$, where x_i is the share of the surplus agent i receives (recall that the time interval

[3] The time interval between two stages is set equal to unity in the simulations.

between two rounds is equal to unity). This model is in fact the finite-horizon variant of Rubinstein's well-known alternating-offers protocol [17].

Two special cases of this model can be identified directly. When $\delta_1 = \delta_2 \to 0$ the entire bargaining surplus vanishes after a single round. This game is therefore equivalent (in terms of payoff) with the abovementioned ultimatum game (see Section 4.1). The special case in which the agents' discount factors approach unity ($\delta_1 = \delta_2 \to 1$) is studied in Section 4.2.

Model with a risk of breakdown (Section 7) The alternating-offers game with a risk of breakdown is examined next. This model is also an extension of the model studied in Section 4.2. However, negotiations do not always proceed to the next stage in this case. Instead, continuation of the bargaining process occurs with a probability $p < 1$. Breakdown of the negotiations may occur in reality when agents get dissatisfied as negotiations take too long, and therefore walk away from the negotiation table, or when intervention of a third party results in a vanishing bargaining surplus. When a negotiation is broken off prematurely both agents receive nothing. In the game-theoretic analysis (see Appendix 1), the continuation probability p plays a similar role as a common discount factor δ ($\equiv \delta_1 = \delta_2$) in the above model with time preferences.

Model with risk averse agents (Section 8) This model extends the previous one by introducing risk averse agents. In this case, agent i's payoff at time t is equal to $u_i(x_i)p^t$, where the degree of risk aversiveness is reflected in the shape of the (concave) utility function u_i. The preferences of the risk averse agents are modeled in this paper with an elementary power-law function, i.e., $u_i(x_i) = x_i^{r_i}$ (with $0 < r_i < 1$ and $i = 1, 2$). The degree of risk aversiveness of agent i is then controlled by the "risk coefficient" r_i.

3. Experimental Setup

We use an evolutionary algorithm (EA) to evolve the negotiation strategies of the agents. Section 3.1 gives an outline of the EA, and discusses how such a system can be interpreted as a model for social or economic learning processes. Our implementation is based on "evolution strategies" (ES), a branch of evolutionary computation that traditionally focusses on real-coded problems [1].[4] The "genetic" representation of the agents' strategies is presented in Section 3.2. The main components of the EA (selection, mutation, and recombination) are discussed in more detail in Sections 3.3 through 3.5.

3.1 The Evolutionary Algorithm

Our evolutionary model consists of two separate populations. Agents in population 1 always start the game (i.e., they are of the "agent 1" type); agents from population 2 are of the "agent 2" type. During a fitness evaluation, an agent plays against a group of opponents who are drawn at random (without replacement) from the other population. The agent's fitness is then equal to the mean utility obtained against these opponents.

This model with two coevolving populations is appropriate if one group of agents has the privilige to open the negotiations. In reality this situation frequently occurs when a potential client wants to buy something from a professional seller. Normally, the seller takes the initiative: he or she can either refer to the indicated price on the product, or propose an initial price.

The EA updates the agents' strategies in successive iterations (also called "generations"). The different stages within one generation are depicted in Fig. 1. First, the fitness of the parental agents is determined by competition between the agents in the two populations. In the next stage (see Fig. 1), "offspring" agents are created. An offspring agent is generated in two steps. First, an agent in the parental population is (randomly, with replacement) selected. This agent's strategy is then mutated to create a new offspring agent (the mutation model is specified in detail in Section 3.4). The fitness of the new offspring is evaluated by interaction with the parental agents. A social or economic interpretation of this parent-offspring interaction is that new agents need to be able to compete with

[4] The widely-used genetic algorithms (GAs) are more tailored toward binary-coded search spaces [10].

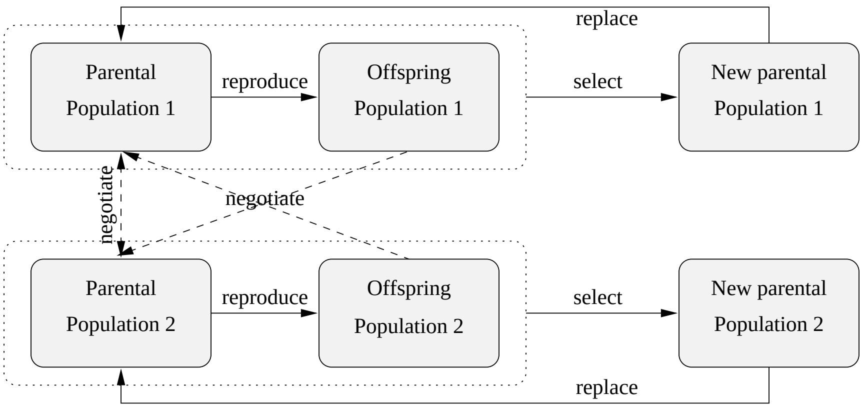

Figure 1: Iteration loop of the evolutionary algorithm (EA). Two populations of agents are evolved separately. Agents in population 1 always start the bargaining process. In the fitness evaluation, both the offspring and the parental agents compete against agents in the two parental populations. The best candidates of the union of parents and offspring are then selected to be the parents in the next iteration.

existing or "proven" strategies before they gain access to a market.[5] In the final stage of the iteration (see Fig. 1), the fittest agents are selected as the new "parents" for the next iteration (see Section 3.3 for more details). This final step completes one iteration of the EA.

All relevant settings of the evolutionary system are listed in Table 1. Pseudo-code of the EA can be found in Appendix 2.

Encoding of chromosome	Real coding
Length of chromosome (l)	n
Mutation	Zero-mean Gaussian ($\sigma = 0.1$)
Recombination	No recombination (see Section 3.5)
Selection	($\mu + \lambda$)-ES
Parent population size (μ)	100
Offspring population size (λ)	100
Number of opponents	25

Table 1: Default settings of the evolutionary model.

3.2 Genetic Representation

Each agent's strategy is encoded as a sequence of real-coded genes (together called a "chromosome"). Assume that agent i from population 1 competes against agent j from population 2. At $t = 0$, agent i then starts the bargaining game and reads the first gene on his chromosome. This "offer" gene, denoted as $o_i(t = 0)$, contains agent i's initial proposal to partition the surplus (i.e., agent i demands a fraction of $o_i(t = 0)$ of the bargaining surplus). Agent j evaluates this offer by inspecting his first gene, which contains the "threshold" value $\tau_j(t = 0)$. Agent j accepts agent i's proposal if $u_j(o_i(t = 0)) \geq \tau_j(t = 0)$. Otherwise, the bargaining process continues. Roles then switch and agent j inspects his second gene, which contains a (counter) offer $o_j(t = 1)$. Agent i evaluates this proposal

[5] In an alternative model, not only the parental agents are used as opponents, but also the newly-formed offspring. This leads to a much more diverse collection of opponents. The fitness of the agents therefore becomes more subject to noise.

by inspecting his second gene, which contains a threshold value $\tau_i(t=1)$. If $u_i(o_j(t=1)) \geq \tau_i(t=1)$, agent i accepts the counter offer. Agent j then receives a share of $o_j(t=1)$ (and agent i the remainder). Otherwise, play continues in an alternating fashion until an agreement occurs or until the deadline is reached (after n rounds).

Note that the length l of each agent's chromosome is equal to n. Because the agents bargain over a surplus of size 1, the offers and thresholds are restricted to the unity interval. The agents' strategies are initialized at the beginning of each EA run by drawing a random number in the unit interval for each gene (from a uniform distribution).

3.3 Selection Scheme

Selection is performed using the $(\mu+\lambda)$-ES selection scheme [1]. In conventional notation, μ is the number of parents and λ is the number of generated offspring ($\mu = \lambda = 100$, see Table 1). The μ survivors with the highest fitness are selected from the union of parental and offspring agents. This selection scheme is therefore an example of an "overlapping generations" model, in which successful agents can survive for multiple generations. A nonoverlapping generations model, in which all parents are discarded after one generation, is investigated in Section 5.1. A probabilistic variant of $(\mu+\lambda)$-ES selection is studied in Section 5.4.

An offspring agent is generated in two steps. First, an agent in the population is (at random, with replacement) selected to be a parent. The chromosome of this parental agent is then mutated to generate a new offspring agent (the mutation model is specified below in Section 3.4). By default the parent-to-offspring ratio is set equal to unity (i.e., $\mu = \lambda$). In Section 5.2 we report several experiments in which this ratio is not equal to unity, to determine the influence of the selection intensity.

In an economic context, selection can be interpreted as imitation of behaviour which seems promising. In general, EAs use two additional operators: mutation and recombination. These operators are explained in detail below.

3.4 Mutation Model

Mutation can be interpreted as undirected exploration of new strategies, or as mistakes made during imitation. In the default mutation model, the offspring's genes x_i are created by adding a zero-mean Gaussian variable, with standard deviation $\sigma_i = 0.1$ [i.e., $N_i(0, 0.1)$],[6] to each corresponding gene x_i of the parent.[7] All offspring genes with a value larger than unity (or smaller than zero) are reset to unity (respectively zero). A mutation model with self-adaptive standard deviations is studied in Section 5.5.

3.5 Recombination Model

Communication between the agents is often modeled by a recombination (or "crossover") operator, which typically exchanges parts of the parental chromosomes to produce new offspring. Recombination of genetic information has proven to be a very effective search operator if the individuals are binary-coded [10]. Following this lead, several recombination models have also been proposed for evolutionary models with real-coded individuals [1].

We performed experiments with two recombination models frequently used in the field of ES: discrete recombination and intermediate recombination [1, pp. 73-78]. However, we did not find a significant change of the fitness of the evolving agents if recombination was allowed (compared to experiments with mutation only). We therefore focus on mutation-based models in this paper.

[6] The notation $N_i(.,.)$ denotes that the random variable is drawn again for each value of the index i.

[7] Notice that the symbol x_i is used in two different meanings. x_i denotes the share received by the i-th agent in bargaining literature and an agent's i-th gene in the field of evolutionary computing. Which usage is appropriate can be inferred easily from the context.

4. Model with Symmetric Time Preferences

All agents have identical discount factors in this section (i.e., $\delta_1 = \delta_2$). Results for $\delta_1 \neq \delta_2$ are presented in Section 6. Unless indicated otherwise, the negotiations are terminated after 10 rounds in our experiments (i.e., $n = 10$). Figure 2 shows the SPE partitioning of the bargaining surplus for a 10-stage game as a function of the common discount factor δ (see Table 2 in Appendix 1). SPE

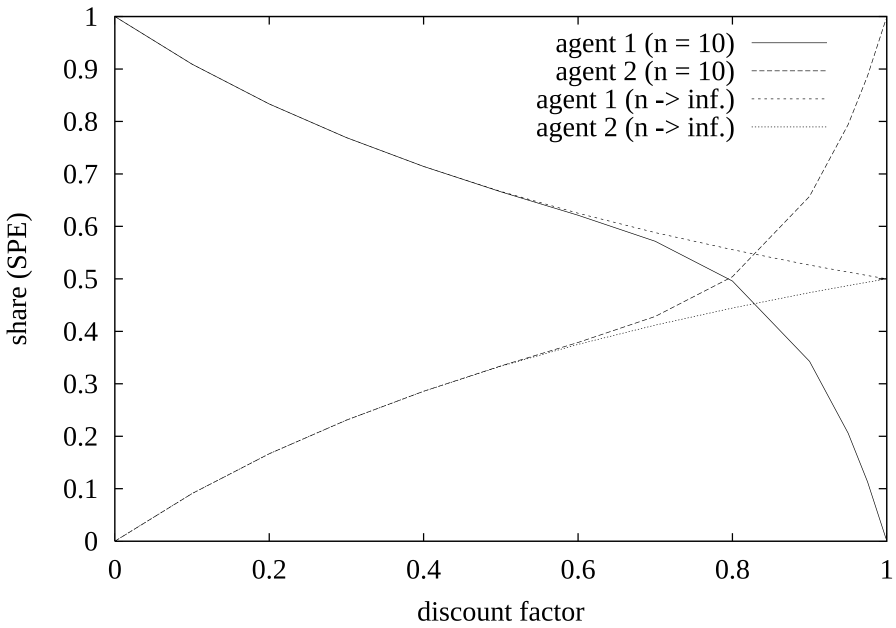

Figure 2: Subgame-perfect partitioning of the bargaining surplus for a 10-stage game. Predictions for an infinite-horizon game are also shown for comparison. Notice the significant differences between these models if the agents' discount factors approach unity.

predictions for the infinite-horizon model are shown as well for comparison. Note that the predictions of the infinite-horizon model only match well with the results of the (exact) finite-horizon model if $\delta < 0.6$ (for $n = 10$). For larger values of δ the discrepancies become significant.

We first investigate two extreme cases. In Section 4.1 we let $\delta \to 0$. Agents then receive nothing if they do not reach agreement in the first round. In payoff terms, this situation is equivalent with the ultimatum game. Another extreme case is obtained by letting $\delta \to 1$. In this case, analyzed in Section 4.2, the agents are payoff-indifferent between reaching a deal sooner or later (provided $t < n$). Figure 2 already shows that the finite horizon of the game becomes important under these conditions. Results for intermediate values of δ $(0 < \delta < 1)$ are summarized in Section 4.3. The influence of changes in our evolutionary model (e.g., different selection schemes, selection intensities, etc.) on the predictions is assessed in Section 5.

4.1 $\delta \to 0$ (The Ultimatum Game)

Figure 3 shows the evolution of the mean fitnesses of the agents in population 1 (the proposers) and population 2 (the responders) in the ultimatum game. Notice that the fitness of agents in population 1 increases rapidly (i.e., the proposers increase their offers for $t = 0$) whereas the fitness of agents in population 2 simultaneously decreases (i.e., the responders lower their thresholds for $t = 0$). Game theory predicts that a rational proposer demands the whole surplus, which the responder accepts (see Appendix 1). This unique (subgame-perfect) equilibrium indeed appears to be an attractor for the evolutionary system.

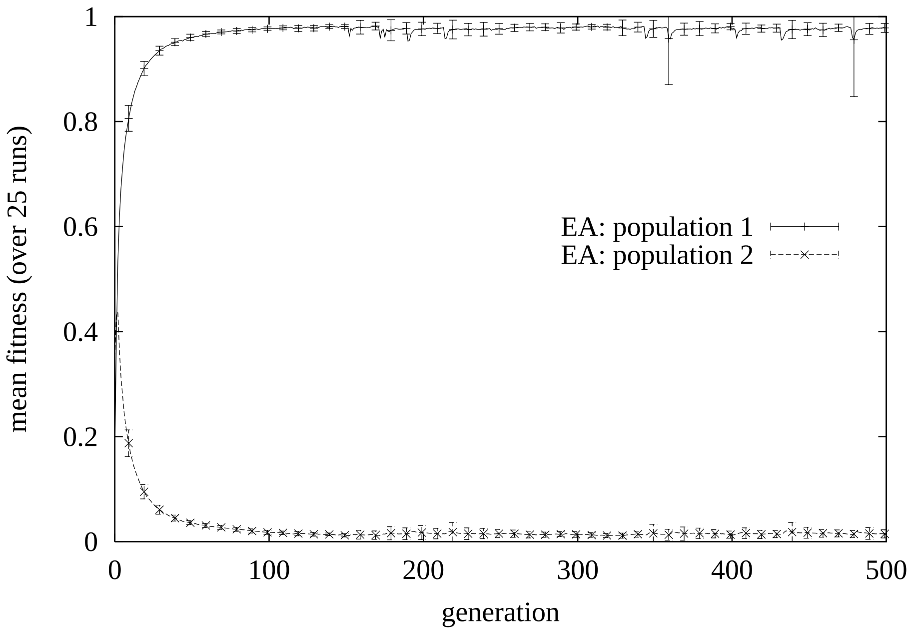

Figure 3: Evolution of the mean fitnesses of agents in population 1 and population 2 in the ultimatum game. Note the clear advantage of agents in population 1 (proposing an offer). Means and standard deviations are calculated across 25 runs.

A closer look at the long-term evolution in Fig. 3 reveals, however, that there is no complete convergence to subgame-perfect behavior. The standard deviations in the fitnesses remain significantly larger than zero, even in the long run, which indicates that the outcomes vary in individual experiments. Figure 4 indeed shows that quite large and sudden transients can occur in individual experiments. The strongly nonlinear behavior visible in Fig. 4 is typical for a "relaxation" oscillation [23, Ch. 12]. Relaxation oscillations are periodic phenomena with very special features during an oscillation. Such an oscillation is typically characterized by intervals of time in which the system changes slowly, followed by short intervals of time in which notable changes occur.[8]

The mechanism causing the oscillations in Fig. 4 is as follows. From an initial start, the agents in population 1 discover that they can increase their offer. The agents in population 2 simultaneously decrease their thresholds (otherwise many disagreements would result, yielding a payoff equal to zero). This process continues until effectively all agents in population 1 demand the whole surplus and all agents in population 2 accept this extreme deal (see Appendix 1). At this point, first reached after ≈ 175 generations in Fig. 4, agents in population 2 become (payoff) indifferent between accepting or refusing (they receive nothing in both cases). The ongoing mutation process then creates offspring agents with a threshold larger than zero. These agents have the same fitness as their all-accepting counterparts in population 2. Therefore, some of them invade population 2. This results in a significant number of disagreements and a sharp drop in fitness for agents in population 1. Consequently, some agents in population 1 decrease their demand in order to stop this process and the fitness of the agents in population 2 increases. Then the race between proposers and responders starts all over again, and the process repeats itself (see Fig. 4).

[8]Relaxation oscillations have also been encountered in coevolving populations of predators and preys in natural ecosystems. A simple mathematical model describing such a system has been proposed by Lotka and Volterra (see [23, Ch. 12] for more details and other examples of relaxation oscillations).

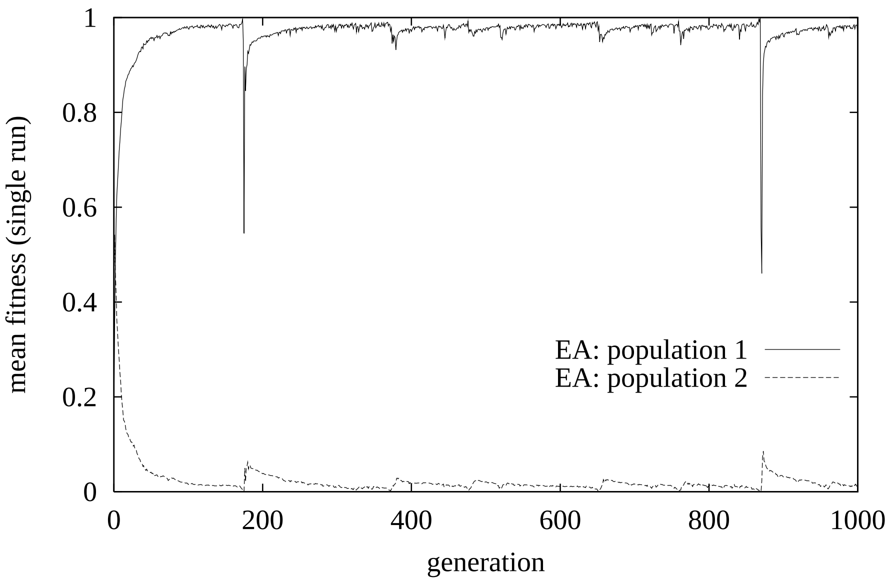

Figure 4: Evolution of the mean population fitnesses in an individual experiment. The strongly nonlinear oscillations occurring in separate runs are masked in the average results shown in Fig. 3. The sharp drops in fitness for agents in population 1 are accompanied by a significant number of disagreements, indicating that agents in population 2 are not accepting the extreme proposals from agents in population 1.

4.2 $\delta \to 1$ (Time Indifference)

Figure 5 shows the evolution of the mean fitness of agents in population 1 in the n-stage alternating-offers game (without payoff discounting). Game theory predicts that the last agent receives the entire bargaining surplus (see Appendix 1). Hence, we would expect the fitness of agents in population 1 to converge to unity if n is odd and to converge to zero if n is even. This tendency is indeed clearly visible in Fig. 5, even for games as long as 20 rounds.

The timing of the agreements is not uniquely defined at subgame perfect equilibrium for this model without time pressure (see Appendix 1). It is therefore of interest to investigate the timing of the agreements in the evolutionary system. We observe that in the experiments most agreements occur just before the deadline in the long run. Consider for instance the 10-stage game. In the first few generations of the evolutionary process, nearly all agreements are reached quickly ($\approx$ 97% of all agreements occur in the first five rounds) and virtually no deals are delayed until the very last round. However, after 25 generations the mean percentage of last-round agreements has already increased to $42 \pm 16\%$. After 500 generations this percentage has increased even further to $80 \pm 3\%$. Interestingly, this deadline-approaching behavior has also been observed in bargaining experiments with humans [16].

Figure 5 furthermore shows that the standard deviations in the mean fitness increase in the long run for short games (i.e., small n). This effect has been observed before in Section 4.1, where it was attributed to unstable population dynamics in the vicinity of the SPE. For multiple-stage games the mechanism triggering these transients is similar (namely, agents who increase their thresholds for $t = n - 1$), but the resulting changes in fitness are relatively small when $n > 1$. Major fitness collapses as detected in the ultimatum game are for instance not occurring in this case.

Partly, stability increases in multiple-stage games (in comparison with the ultimatum game studied

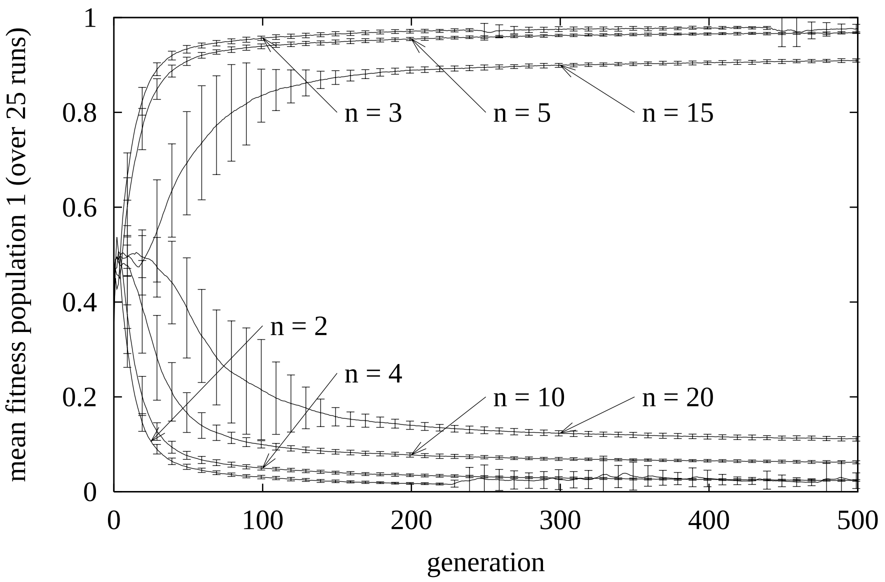

Figure 5: Evolution of the mean fitness of agents in population 1 in n-stage alternating-offers games without payoff discounting. When the agents in population 1 have the opportunity to make the last offer (i.e., when n is odd) they receive the largest share of the surplus in the long run. Exactly the opposite happens when agents in population 2 are the last in turn (i.e., when n is even).

in Section 4.1) because convergence towards extreme SPE behavior takes longer for larger n (see Fig. 5). Recall that agents with a non-zero threshold for $t = n - 1$ can only invade the population if the SPE is actually reached (see Section 4.1). The computer experiments also show that when the SPE is reached temporarily more agreements occur in earlier rounds. The timing of the agreements changes because the proposers in the last round decrease their offers and thresholds in earlier rounds in order to avoid the occurrence of disagreements in the very last round. In the ultimatum game the proposers do not have this opportunity. This leaves them with just one tactic to avoid the occurrence of disagreements (i.e., rapidly increasing the amount they offer to their opponents).

4.3 $0 < \delta < 1$

Figure 6 shows that the fitnesses of the two populations converge very rapidly to a steady state if $\delta = 0.6$. Game theory predicts that the agents agree immediately on a deal if the agents are under time pressure (see Appendix 1). The computational experiments indeed show that the agents hasten to reach a deal. Initially, $51 \pm 4\%$ of all agreements are reached in the first round. After 25 generations this percentage has already increased to $91 \pm 8\%$. After 500 generations this percentage is essentially unchanged ($91 \pm 5\%$). Note that we encountered exactly the opposite behavior (namely, the postponement of agreements) for $\delta = 1$ in Section 4.2.

This suggests that the speed of convergence is affected significantly by the initialization procedure of the agents' strategies at the beginning of each EA run. As noted before in Section 4.2, the random initialization of the chromosomes gives rise to many agreements in the first few rounds. The initial timing of agreements is therefore already very close to the optimal situation for $\delta < 1$ (all Pareto-efficient deals are made in the first round). Hence, the evolving agents quickly fine-tune their strategies and after a short period of time almost all agreements are indeed reached immediately. On the other hand, if $\delta = 1$ the random initialization leads to a far from optimal timing (because initially almost

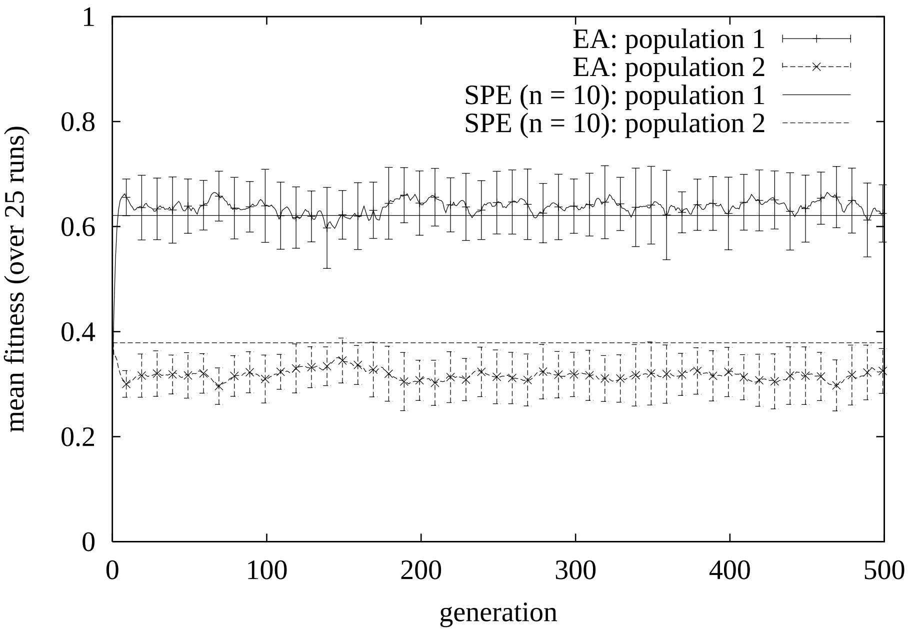

Figure 6: Evolution of the mean fitness for both populations in the 10-stage alternating-offers game (with $\delta = 0.6$). A very rapid convergence of the fitnesses is visible. For comparison, the SPE partitioning for $n = 10$ is also indicated.

no agreements are reached in the last round). The evolving system therefore converges slowly to a steady state [compare the slow transient in Fig. 5 (for $n = 10$) with the rapid convergence in Fig. 6].

The SPE partitioning of the surplus is also indicated in Fig. 6 for comparison. Note that the SPE prediction accurately predicts the share of the surplus that agents in population 1 receive in our experiments. The evolving agents in population 2 receive less than predicted by game theory, however. The total share divided by the evolving agents is smaller than unity in the experiments, because some agreements ($\approx 10\%$ of all deals) are not reached in the first round. Taking this effect into account, it becomes clear from Fig. 6 that the agents in population 1 receive a relatively large share of the (partially discounted) surplus compared to the agents in population 2.

We study the partitioning of the surplus for a wider range of discount factors in Figs. 7 and 8. These figures again show that agents in population 1 receive more than game theory predicts (see Fig. 7), whereas agents in population 2 negotiate relatively poor deals (see Fig. 8). This effect is particularly clear in case of strong time preferences (for instance when $\delta \approx 0.3$). An explanation might be that rejecting offers in the first round has a strong negative effect on the fitness of agents in population 2 if δ is small. Agents in population 2 will therefore strongly prefer to reach an agreement in the first round. This tendency can subsequently be exploited by agents in population 1 by demanding a relatively large share of the surplus (as in the ultimatum game discussed in Section 4.1). This hypothesis is supported by the experimental observation that almost all agreements are reached immediately (after 500 generations, more than 98% of all agreements are reached in the first round for $\delta = 0.3$). This means that in almost all cases a very short game is played (only one stage).

Figures 7 and 8 furthermore show that in case of weak time pressure (for instance when $\delta \approx 0.9$) the bargaining outcome deviates significantly from the SPE prediction for $n = 10$. Figure 8 for instance shows that the agents in population 2 do not fully exploit their last-mover advantage under these circumstances (their mean fitness is far below the SPE level). This effect can be explained by the boundedly-rational behavior of the adaptive agents. These agents do not reason backwards from the

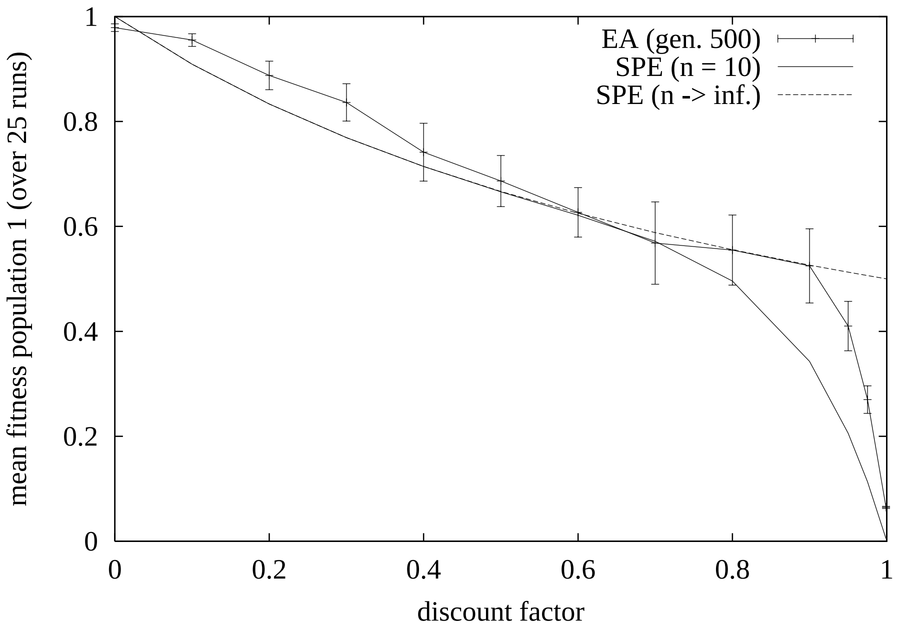

Figure 7: Performance of agents in population 1 as a function of the discount factor. Game theoretic predictions for the 10-stage game and the infinite-horizon game are also shown for comparison. Notice the significant deviations from these predictions if the agents have either weak or strong time preferences.

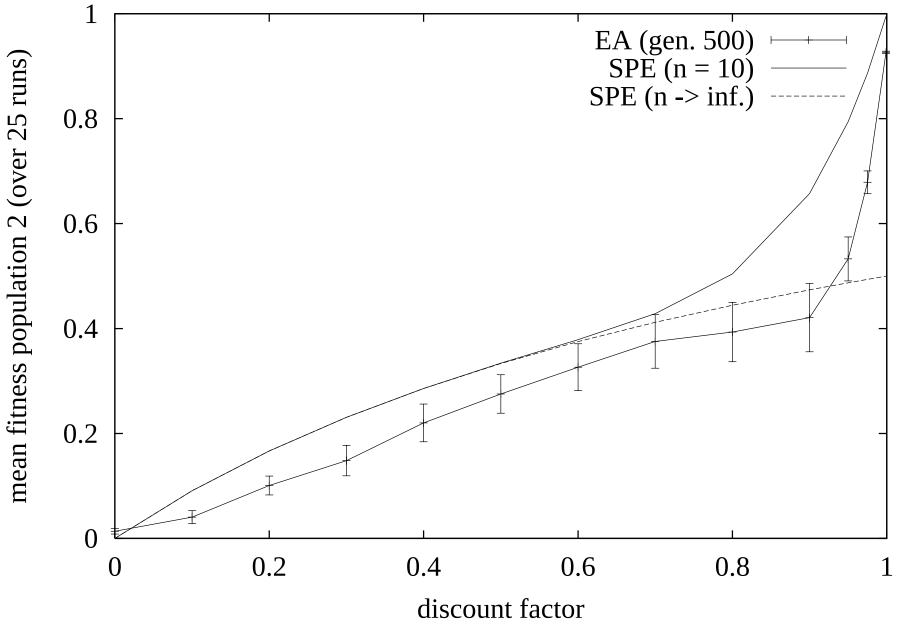

Figure 8: Performance of agents in population 2 as a function of the discount factor. Notice that the agents in population 2 negotiate poor deals [relative to game-theoretic (SPE) predictions].

deadline (as is done in game theory, see Appendix 1), but focus on the first few rounds, where expected utility is relatively high. This means that only few agreements are reached in later rounds. As a result, the deadline of the game is not perceived accurately by the evolving agents.

In fact, the experimental results agree much better with SPE predictions for longer games. Almost perfect agreement is for instance obtained (for large δ) if we compare the experimental results with SPE predictions for a 30-stage game. This lends more support to Rubinstein's analysis of an infinite-horizon game: in reality an infinite game length may be a good modeling assumption if the agents do not perceive the finite deadline of the game. Figures 7 and 8 indeed show that the experimental outcome is predicted quite well (for δ up to 0.9) by theoretical predictions for an infinite-horizon game.

5. Alternative Evolutionary Models

The specific choice of settings for an evolutionary model can affect the obtained results substantially. A telling example has recently been given for the evolutionary prisoner's dilemma [21]. In particular, it was shown in [21] that the course of evolution in a multi-agent system can be very sensitive to modeling choices like (i) whether successive generations of agents overlap or not, or (ii) the selection intensity in the population. The influence of these two factors will therefore be investigated in Sections 5.1 and 5.2. In addition, we evaluate the influence of the population size in Section 5.3. Section 5.4 investigates an alternative selection model, in which selection is probabilistic instead of deterministic. Finally, Section 5.5 investigates a model in which the agents can adapt their own mutation step-sizes.

Figure 9 summarizes results for the model variants considered in the remainder of this section. This

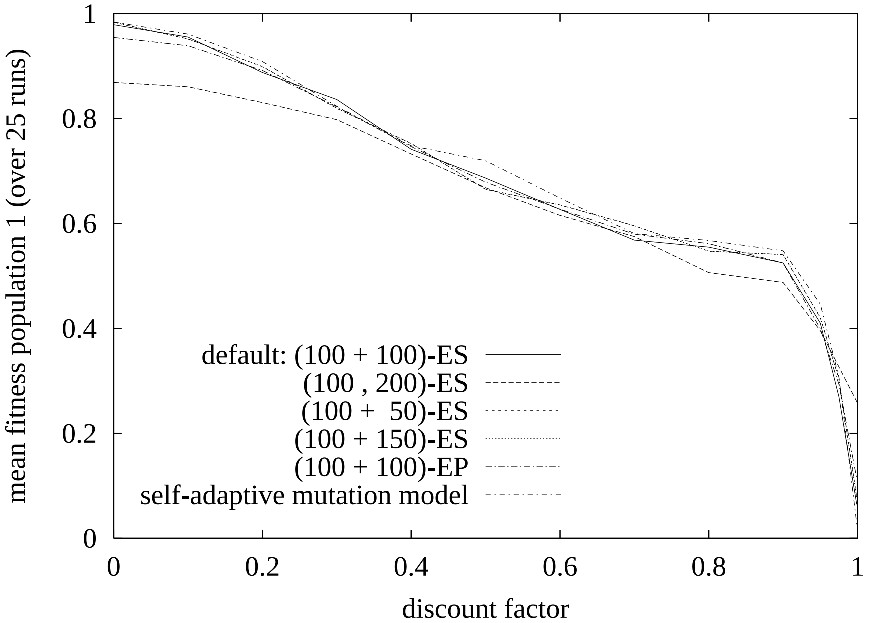

Figure 9: Long-term fitness of agents in population 1 as predicted by different evolutionary models (for $n = 10$). Apart from the $(100, 200)$-ES model, all models yield rather similar results. Note that deviations for the $(100, 200)$-ES model become especially large for $\delta \to 0$ and $\delta \to 1$. (For clarity, standard deviations are not shown.)

figure shows the long-term fitness of agents in population 1 (measured after the initial transients have died out). Predictions of the default selection scheme [the $(100 + 100)$-ES model, see Table 1] are indicated with a solid line.

5.1 Overlapping vs. Nonoverlapping Generations Models

As we mentioned before in Section 3, the $(\mu + \lambda)$-ES selection scheme is an "overlapping generations" model in the sense that well-performing agents can survive for extended periods of time. As an alternative, we also consider a "nonoverlapping generations" selection scheme proposed in the field of ES. In this (μ, λ)-ES model [1] μ parents produce $\lambda > \mu$ offspring. All parents are discarded after one period and only the μ best offspring are transferred to the next generation. Unless indicated otherwise, we set λ equal to $2\mu = 200$ in this model. Hence, the number of agents competing for survival is the same as in the experiments with the default $(100 + 100)$-ES model.

Figure 9 shows that predictions of the $(100, 200)$-ES model differ significantly from those obtained with a $(100 + 100)$-ES scheme. Differences become especially large for $\delta \to 0$ and $\delta \to 1$. Under these circumstances, the agents basically play the ultimatum game: either in the first round (when $\delta \to 0$) or in the last round (when $\delta \to 1$).

As an example, we investigate the limiting case $\delta \to 0$ in more detail. In this situation, game theory predicts (see Appendix 1) that the agents in population 1 demand the whole bargaining surplus in the first round (i.e., at $t = 0$). This behavior is not observed in the evolutionary experiments: the offer gene $o(t = 0)$ of the agents in population 1 evolves to a value of 0.90 ± 0.04.[9] The threshold gene $\tau(t = 0)$ of the agents in population 2 evolves to a value close to zero (≈ 0.01), in good agreement with the SPE prediction that these responding agents accept all proposals (see Appendix 1). As a consequence, the mean fitness of the agents in population 1 [which is ≈ 0.87 for $\delta \to 0$, see Fig. 9] is approximately equal to the mean offer they submit.

The deviation from SPE behavior by the agents in population 1 is specific for the nonoverlapping generations model under consideration here. In such a model, only the offspring are transferred to the next generation. Because all offspring have just been mutated, a significant spread occurs in the proposals submitted by these agents. In ultimatum game situations, the mean offer of the proposing agents is therefore significantly lower than the extreme offer of 1.0 predicted by game theory. As a result, the fitness of the proposing agents converges to a level below unity.[10]

Based on the above discussion one may expect that the agents' behavior agrees better with SPE predictions if their mutation step sizes σ_i become smaller. If we set σ_i equal to 0.025 (in the default model $\sigma_i = 0.1$, see Table 1), $o(t = 0)$ converges to a value of 0.95 ± 0.01. As a consequence, the fitness of the agents in population 1 increases to ≈ 0.95. This value is already much closer to the SPE prediction of 1.0.

Increasing the selection intensity should have a similar effect, because selection discards inferior offspring (such as offspring in population 1 submitting low offers). If we increase the selection intensity by setting λ equal to 4μ (instead of 2μ), $o(t = 0)$ converges to a value of 0.96 ± 0.02. This leads to a long-term fitness of ≈ 0.93 for the agents in population 1. When λ/μ becomes large (for instance when $\lambda/\mu = 7$, a common setting for the (μ, λ)-ES model [1]) the evolving system can approach the SPE. Under these circumstances relaxation oscillations occur, for reasons pointed out before in Section 4.1.

Note that the assumption that successive generations of agents do not overlap is essential to explain the convergence to equilibria which are not subgame-perfect in the ultimatum game. In overlapping generations models only well-performing offspring can enter the populations of proposers and responders.[11] This means that offspring submitting low offers will not gain access to the next generation. The same holds for offspring with a relatively high threshold. These will not gain access to the population of responders (except when the SPE is actually reached), because such agents cause relatively many disagreements and therefore earn a lower payoff than their parents.

[9] The standard deviation is a measure of how widely the offers (at $t = 0$) of the agents in population 1 are dispersed from the mean.

[10] Convergence to an equilibrium which is not subgame-perfect in the ultimatum game has been reported before in the field of evolutionary game theory [7]. In [7] the evolution of strategies for the ultimatum game is governed by a nonoverlapping generations variant of the replicator dynamics [2, Ch. 9].

[11] Assuming that selection is strictly deterministic. See Section 5.4 for a further discussion.

5.2 Influence Selection Intensity

We now further investigate the (default) ($\mu+\lambda$)-ES model. Figure 9 shows that decreasing the selection intensity in this model, by setting λ equal to 50, does not affect the mean results (in comparison with the default model in which $\lambda = 100$, see Table 1). The same conclusion holds if we increase the selection intensity by setting λ equal to 150. The mean results in Fig. 9 may, however, mask sudden collapses in fitness when $\delta \to 0$, as we noticed before in Section 4.1. Individual experiments for $\delta \to 0$ show that these collapses occur more frequently if λ is increased to 150. Unstable behavior disappears almost completely, on the other hand, if $\lambda = 50$. Inspection of individual runs shows that major collapses in fitness only occur if *all* agents adopt subgame-perfect behavior (see Fig. 4). This explains the impact of the selection intensity on stability: a population is more likely to converge to a uniform state if the selection intensity becomes large.

5.3 Influence Population Size

The above discussion suggests that evolutionary stability might deteriorate for $\delta \to 0$ in undersized populations, because in this case genetic diversity is rapidly lost as a result of natural selection. This effect is indeed observed in additional experiments. Small populations converge more quickly toward SPE behavior initially, but are very unstable in the long run. The fitnesses of the coevolving populations even start to oscillate with a period as short as 25 generations if the population size is reduced to 25 (and $\delta \to 0$). This oscillation period becomes much larger ($\approx$ 5000 generations) if the population size is increased to 200. Relaxation oscillations are therefore most likely to occur in small populations of adaptive agents.

5.4 Deterministic vs. Probabilistic Selection

One may argue that selection in economic markets is often of a probabilistic nature. The assumption in deterministic models that the worst agents are never selected ("imitated") can for instance be considered as an idealization: in reality imperfect selection occurs in actual market situations. Selection of the agents may also be based on a limited number of comparisons with other competitors. Such a "tournament setting" introduces additional uncertainty in the selection process. Inferior agents can for instance survive if they are paired with even weaker competitors, whereas relatively fit agents can be eliminated in a tournament with top-performers.

We investigate the influence of these stochastic processes by considering a probabilistic variant of the ($\mu+\lambda$)-ES selection scheme. This alternative scheme has been proposed in the field of evolutionary programming (EP) [1, pp. 96-99]. In this model, denoted here as ($\mu + \mu$)-EP selection, the parental and offspring populations have an equal size μ (i.e., $\lambda = \mu$). Each agent is evaluated against q agents from the union of parents and offspring. These opponents are selected at random (with replacement). A typical setting for the tournament size q is 10 (or smaller) [1, p. 102]. If an agent's fitness is larger than (or equal to) his opponent's fitness, he is attributed a "win". The μ agents with the largest number of wins are transferred to the next generation. It has been shown in [1, p. 96-99] that this tournament scheme selects the same agents as the deterministic ($\mu + \mu$)-ES if $q \to \infty$.

Figure 9 shows that predictions for the ($100 + 100$)-EP model (with $q = 10$) are in first approximation rather similar to results obtained with the deterministic ($100 + 100$)-ES model. Slight deviations occur, however, for $\delta \to 0$ and $\delta \to 1$. To explain these differences, we study the situation in which $\delta \to 0$ in more detail. We observed earlier in Section 4.1 that under these (ultimatum game) conditions the evolving system can actually reach the SPE in individual runs. This state was subsequently undermined by responding agents with nonzero thresholds. In case of probabilistic selection this destabilizing process might actually start *before* the SPE is reached. An additional effect of probabilistic selection is, however, that the selective pressure decreases. This has a stabilizing effect, as we pointed out in Section 5.2.

Inspection of individual runs (for $\delta \to 0$) shows that the latter (stabilizing) effect is dominant. The evolving system converges to a rather stable state which is close to, but removed from, the SPE. We can therefore conclude that relaxation oscillations are less likely to occur when selection is stochastic instead of deterministic.

More extreme cases (viz., $q = 2$ or $q = 100$) have also been studied in additional experiments. As expected, results for $q = 100$ are very similar to predictions of the deterministic $(100 + 100)$-ES. If $q = 2$, significant deviations from SPE predictions occur (especially if $\delta \to 0$ or $\delta \to 1$). The mean long-term fitness of the agents in population 1 is for instance equal to 0.83 ± 0.01 for $\delta \to 0$, significantly below the SPE prediction of 1.0. Inferior agents in population 1 remain in the population with a relatively large probability in this case. This leads to a spread in the offers submitted by these agents in ultimatum game situations. Their mean offer therefore converges to a value below the SPE level (follow the similar discussion in Section 5.1).

5.5 Fixed vs. Self-Adaptive Mutation Step-Sizes

The default mutation model studied in this paper, see Section 3.4, sets the standard deviations of all agents to a common value. This value then remains constant during the course of evolution. A more natural approach would be to enable individual agents to control the magnitude of the mutations in their genetic code. An elegant mutation model which can be used for this purpose has been described in [1, pp. 71-73]. This model allows an evolutionary self-adaptation of both the genes and the corresponding standard deviations at the same time. More formally, an agent consists of object variables $[x_0, ..., x_{l-1}]$ and strategy parameters $[\sigma_0, ..., \sigma_{l-1}]$ in this model.

The mutation operator first updates an agent's strategy parameters σ_i into σ_i'-values in the following way:

$$\sigma_i' := \sigma_i \exp[\tau' N(0,1) + \tau N_i(0,1)], \tag{5.1}$$

where τ' and τ are the so-called "global" and "individual" learning rates. We use commonly recommended settings for these parameters.[12] After the strategy parameters have been modified, the object variables are mutated: $x_i' := x_i + \sigma_i' N_i(0,1)$.

The initial standard deviations $\sigma_i(0)$ are set to a value of 0.1 (the same value as in the default mutation model) to close this mutation model. The particular value chosen for $\sigma_i(0)$ is not expected to be crucial, because the self-adaptation process rapidly scales the step sizes into the proper range. To prevent complete convergence of the population, we force all standard deviations to remain larger than a small value $\varepsilon_\sigma = 0.025$ [1, pp. 72–73]. Note that this self-adaptive model reduces to the default mutation model with fixed standard deviations if we set the learning rates τ' and τ equal to zero.

We observe in the experiments with this self-adaptive mutation model that the agents in general reduce their mutation step-sizes in the course of evolution (i.e., their search in the strategy space becomes more local instead of global). Figure 9 nevertheless shows that the prediction of the mean fitness is very similar to the prediction of the model without step-size control. This indicates that the $(\mu + \lambda)$-ES selection scheme successfully discards less successful mutants in the mutation model with constant step-sizes.

We demonstrated in Section 5.1 that the proliferation of inferior offspring leads to significant deviations from SPE behavior in a (μ, λ)-ES model. It can be expected that these discrepancies become smaller in such a model if the agents are able to reduce the magnitude of the mutations and produce more similar offspring. This is indeed the case. If we use a $(100, 200)$-ES and let $\delta \to 0$, the mean long-term fitness of the agents in population 1 is ≈ 0.93 (compared to ≈ 0.87 in the experiments with fixed mutation step-sizes, see Section 5.1). This value is indeed in better agreement with the SPE prediction of 1.0.

6. Model with Asymmetric Time Preferences

We now continue our experiments with the default evolutionary model specified in Table 1. Figure 10 shows the long-term performance of the evolving agents in case of asymmetric time preferences. For

[12] Namely, $\tau' = (\sqrt{2l})^{-1}$ and $\tau = (\sqrt{2\sqrt{l}})^{-1}$ [1, p. 72].

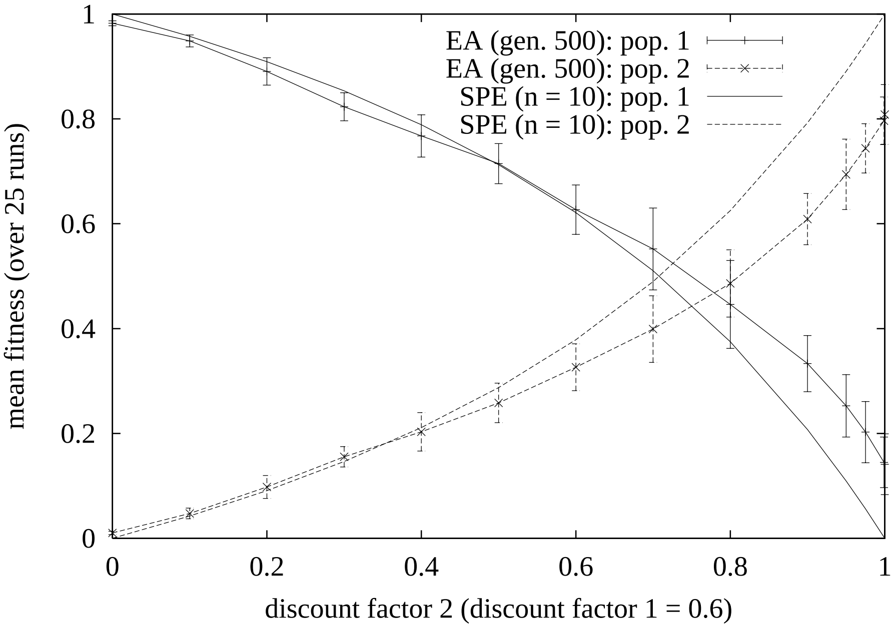

Figure 10: Long-term fitnesses of agents in population 1 and population 2 in case of asymmetric time preferences. The discount factor for agents in population 2 (δ_2) is varied between zero and unity in this figure. Agents in population 1 have a fixed discount factor ($\delta_1 = 0.6$). Notice that agents in population 2 do not fully exploit their bargaining power (in particular when $\delta_2 > \delta_1$).

comparison, the game-theoretic predictions are shown as well.[13] All agents in population 1 have a fixed discount factor δ_1 (equal to 0.6) while the discount factor of agents in population 2 (δ_2) is varied between zero and unity. The fitnesses of the agents converge very rapidly (within 50-150 generations) to the values reported in Fig. 10. Note that the performance of the agents in population 2 is not as good as predicted by game theory when $\delta_2 > \delta_1$, whereas the agents in population 1 actually do better. This effect becomes especially clear if $\delta_2 \to 1$. We will study this case in more detail below.

When $\delta_2 \to 1$, the agents in population 2 experience no time pressure to reach an early agreement. Time pressure is, on the other hand, relatively large for the agents in population 1 (their payoff dimishes proportional to 0.6^t as a function of the round number t). The agents in population 1 thus become more and more (payoff) indifferent as the bargaining game proceeds. This reduces the evolutionary pressure against agents in population 1 with large values for their threshold and offer genes (for large t). In the experiments we even observe that these genes evolve to random values (in the unit interval) for $t \geq 4$.

Agents in population 2 who try to exploit their bargaining power by delaying agreements will therefore encounter opponents with random strategies in later rounds. This deprives these agents partly of their bargaining power: they cannot force their indifferent opponents to adjust their behavior in later rounds. In fact, exactly the opposite occurs in the evolutionary system. In an attempt to avoid the occurrence of disagreements, the agents in population 2 reduce their offers and thresholds in later rounds.[14]

Experiments with the alternative models discussed in Section 5 lead to similar results for $\delta_2 > \delta_1$.

---

[13]The SPE partitioning for an infinite-horizon game is not indicated in Fig. 10. Differences with the prediction for the 10-stage game are very small (at most of the order of 0.01 for $\delta_2 \approx 0.85$).

[14]$\tau(t=0)$ evolves for instance to 0.78±0.14 in population 2, whereas $\tau(t=4) = 0.64\pm0.26$ and $\tau(t=8) = 0.32\pm0.25$ in the long run. The offer genes evolve to values of $o(t=1) = 0.90\pm0.09$, $o(t=5) = 0.76\pm0.32$, and $o(t=9) = 0.55\pm0.31$.

Hence, the deviations from game-theoretic predictions in the computational experiments cannot be attributed to the specific settings of the $(100 + 100)$-ES (which was used to generate Fig. 10).

7. Model with a Risk of Breakdown

Our bargaining model with stochastic breakdown generates exactly the same game-theoretic solutions as the model studied in Section 4 if the continuation probability p is set equal to δ (see Appendix 1). It is therefore instructive to compare the experimental results in Fig. 11 with those reported previously in Figs. 7 and 8. This comparison shows that the long-term behavior of the evolving agents is different for

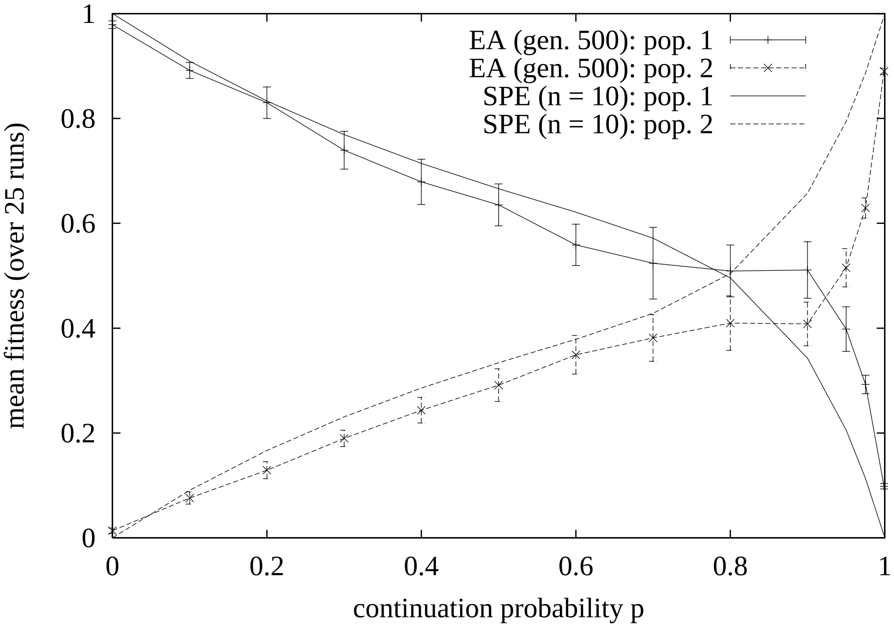

Figure 11: Performance of agents in population 1 and population 2 as a function of the continuation probability p. Note that the experimental trends are predicted very accurately by game theory if p is smaller than 0.7. Deviations become significant, however, for larger continuation probabilities.

small values of p and δ. Notice for instance that the experimental trends in Fig. 11 correspond rather well with game-theoretic predictions for $p < 0.7$, whereas we observed more significant deviations for $\delta < 0.7$ in Figs. 7 and 8. This improved agreement between experiment and theory is due to the stochastic nature of bargaining with a risk of breakdown. The additional randomization leads to a relatively large diversity of strategies in the early stages of evolution. A premature convergence to “ultimatum game” behavior (as we observed in Figs. 7 and 8 for small values of δ) does therefore not occur for small values of p. The significant deviations from SPE predictions for large δ (see Figs. 7 and 8) do also occur in a model with breakdown as p becomes large (see Fig. 11).

Additional experiments with alternative overlapping generations models (e.g., with a different selection intensity, or with a self-adaptive mutation model) lead to very similar results as those reported in Fig. 11. Simulations with the nonoverlapping generations $(100, 200)$-ES model show, on the other hand, that significant deviations from the SPE predictions occur if p is either close to zero or close to unity. This is not surprising, however, because letting $p \to 0$ in the model with breakdown yields exactly the same bargaining model (viz., the ultimatum game) as by letting $\delta \to 0$ in the model with time preferences. In a similar fashion we can identify $p \to 1$ with $\delta \to 1$. We already explained in Section 5.1 why the evolutionary system can converge to equilibria that are not subgame-perfect under these extreme conditions.

8. Model with Risk Averse Agents

Figure 12 shows the effect of risk aversiveness on the respective shares received in a 10-stage game with breakdown (with $p = 0.6$). In Fig. 12 the risk coefficient of one of the agents is varied between zero

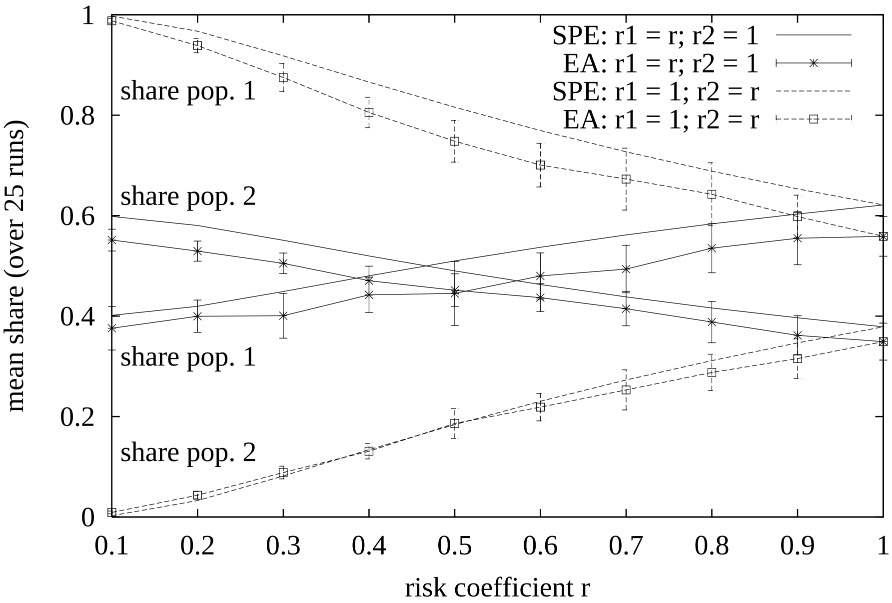

Figure 12: Partitioning of the surplus as a function of the degree of risk aversiveness of both agents (for $n = 10$ and $p = 0.6$). For each curve the risk coefficient of one of the agents is varied between 0.1 and 1.0, whereas his opponent remains risk neutral. Notice that an agent's share increases if his opponent becomes more risk averse. Also, this (positive or negative) effect is not equal for both agents. Notice that the experimental results closely follow the trends predicted by game theory. (EA statistics obtained after 500 generations.)

and unity, while the other agent remains risk neutral (or, equivalent, whose risk coefficient remains equal to unity). The solid line in Fig. 12 for instance indicates that agent 1 receives a smaller share of the surplus if he becomes more risk averse (while agent 2 remains risk neutral). On the other hand, agent 1's share increases if agent 2 becomes more risk averse (while agent 1 remains risk neutral). This effect of risk aversion has been noticed before in literature on bargaining [11].

We also like to point out another, quite subtle, effect. Figure 12 shows that the (SPE) share received by agent 1 increases from 0.625 to ≈ 1.0 if the risk coefficient of agent 2 decreases from 1.0 to 0.1. On the other hand, agent 1's share only slightly decreases (from 0.625 to ≈ 0.4) if his own risk coefficient is reduced from 1.0 to 0.1. Exactly the opposite effect is visible for agent 2. This indicates that if agent 1 becomes more risk averse, this has a relatively small effect on his share, whereas if agent 2 becomes more risk averse, this agent rapidly loses share to his opponent. Figure 12 shows that this subtle effect is also detected by our evolutionary system. The trends predicted by game theory are in fact followed very closely for a wide range of risk coefficients.

These results indicate that it is feasible to investigate the influence of an agent's attitudes toward risk in an evolutionary setting. The evolutionary system correctly predicts that an agent's share of the bargaining surplus diminishes if he becomes more risk averse. We also identified a more subtle second-order effect, namely that the impact of risk aversion on an agent's negotiated share is depending on the bargaining order (i.e., whether the agent can start the negotiations or not).

9. Conclusions

We study equilibrium selection in evolutionary bargaining models. Computational experiments are performed using evolutionary algorithms (EAs). Negotiations between the adaptive agents are governed by a finite-horizon version of Rubinstein's well-known "alternating-offers" protocol. Besides Rubinstein's standard model (with exponential payoff discounting) an alternative model with stochastic breakdown in negotiations is examined. Using this model variant, the influence of risk aversion on the behavior of adaptive agents is assessed.

This paper shows that game-theoretic approaches are very useful to interpret equilibrium-selecting behavior in evolutionary systems of adaptive bargaining agents. The adaptive agents are boundedly rational because they only experience the profit of their interactions with other agents. Nevertheless, they display behavior that is surprisingly "rational" and fully informed in many instances. Agreement between theory and experiment is especially good when the agents experience an intermediate time pressure.

In extreme situations (i.e., when time pressure becomes either extremely strong or negligible) more significant deviations from game-theoretic predictions emerge. We demonstrate that the specific evolutionary model governing agent selection is an important factor in this case. In "overlapping generations" models, the evolutionary system rapidly converges to subgame-perfect behavior (but highly nonlinear transients can occur in this case). In "nonoverlapping generations" models, on the other hand, convergence to equilibria which are not subgame-perfect is observed.

Two other experimental observations should be mentioned here. First, the finite horizon of the negotiations is not always fully exploited by the last agent in turn (even if time pressure is rather weak). In fact, the boundedly-rational agents often act as if the length of the game is actually much longer. This lends more support to the "infinite-horizon" assumption frequently employed in game-theoretic work. This approximation may yield surprisingly accurate results in evolutionary systems when the agents do not perceive the deadline of the negotiations. Second, we observe (and explain) discrepancies between theory and experiment if the agents are asymmetric (i.e., when they discount the bargaining surplus at different rates).

More in general, this work presents a systematic validation of evolutionary and computational techniques in the field of bargaining. As a promising line of research we are currently studying adaptive agents with more complex bargaining strategies. Another important aspect of electronic trading, negotiations over multiple issues, is discussed in a companion paper [8].

Appendix

1. Alternating-Offers Games: Analysis

We will first analyse the one-stage ultimatum game. Subgame-perfect equilibrium strategies for multiple-stage games (with complete information) are then derived by using a backward-induction approach [22, Ch. 1]. It is assumed in this appendix (as has been done throughout this paper) that the size of the bargaining surplus is equal to unity at $t = 0$.

The ultimatum game (Section 4.1) In the ultimatum game, there is a unique SPE in which the proposer (i.e., agent 1) demands 1, and the responder (i.e., agent 2) accepts this offer (and therefore receives nothing). Note that there does *not* exist a SPE in which agent 2 rejects this extreme proposal. A clear justification of this (at first sight somewhat counterintuitive) result can be given by analyzing a finite version of the ultimatum game, in which offers should be made in multiples of a smallest monetary unit ϵ [2, pp. 200–201]. In this game there actually exist two SPEs (viz., a SPE in which agent 1 demands 1 and agent 2 accepts all proposals, and a second SPE in which agent 1 demands $1 - \epsilon$ and agent 2 accepts all demands except 1). However, these two solutions converge to the single SPE of the limiting game with continuous offers if ϵ becomes very small.

Multiple-stage games (Section 4.2) Both agents are indifferent between accepting a deal sooner or later in this model. Hence, the last agent in turn has the opportunity to reject all proposals from his opponent and demand the entire surplus in the last round (which the other agent then accepts). If the maximum number of rounds n is odd, agent 1 will therefore receive the entire surplus, whereas agent 2 receives all in case n is even. Due to the absence of time pressure multiple subgame perfect equilibria exist in this case. Although these equilibria differ in the timing of the agreements, they all result in the same outcome (i.e., the last agent in turn always receives the entire surplus). It is for instance subgame perfect for the last responder to concede the entire surplus to his opponent before the deadline is actually reached or, alternatively, to accept a take-it-or-leave-it deal from the opponent at any point in time.

Model with time preferences (Sections 4.3 through 6) Subgame-perfect equilibrium strategies for the finite-horizon models considered in Sections 4.3 through 8 are derived below by applying the following basic principle [2, pp. 199–200]:

> In (subgame-perfect) equilibrium, a proposer always plans to offer the responder an amount that will make the responder indifferent between accepting and refusing. In equilibrium, the responder always plans to accept such an offer or better, and to refuse anything worse.

Note that applying this principle to the ultimatum game leads to the same result as given above.

Using this principle we calculate the (unique) SPE for a model with time preferences as follows. If the maximum number of rounds n is even, agent 2 will be the proposer in the last round (i.e., at $t = n - 1$). Agent 2 will then demand the whole surplus (of size 1 at the beginning of play, but only of size δ_2^{n-1} in the last round) and agent 1 will receive nothing. This division of the surplus would yield agent 2 a payoff which is equal to δ_2^{n-1}. In equilibrium, at $t = n - 2$ agent 1 should propose agent 2 a payoff-equivalent deal. This implies that agent 1 requests a fraction $1 - \delta_2$ at $t = n - 2$. This division of the surplus would yield agent 1 a payoff equal to $\delta_1^{n-2}(1 - \delta_2)$. This procedure is then repeated until the beginning of the game is reached (at $t = 0$). The same line of reasoning holds if the number of rounds is odd (simply switch the roles of agent 1 and agent 2). In equilibrium, agent 1 then demands a share of $x_1^*(n)$ in the first round and agent 2 immediately accepts this proposal [receiving $x_2^*(n) = 1 - x_1^*(n)$].

The SPE partitioning (x_1^*, x_2^*) as a function of the game length is listed in Table 2. To be expected, this partitioning of the surplus converges to the partitioning derived by Rubinstein for the infinite-horizon game [17]. In Rubinstein's model agent 1 receives $\frac{1-\delta_2}{1-\delta_1\delta_2}$ and agent 2 receives the remaining part of the surplus.

n	SPE share of agent 1 (x_1^*)	SPE share of agent 2 (x_2^*)
1	1	0
2	$1-\delta_2$	δ_2
3	$1-\delta_2(1-\delta_1)$	$\delta_2(1-\delta_1)$
4	$1-\delta_2(1-\delta_1(1-\delta_2))$	$\delta_2(1-\delta_1(1-\delta_2))$
5	$1-\delta_2(1-\delta_1(1-\delta_2(1-\delta_1)))$	$\delta_2(1-\delta_1(1-\delta_2(1-\delta_1)))$
6	$1-\delta_2(1-\delta_1(1-\delta_2(1-\delta_1(1-\delta_2))))$	$\delta_2(1-\delta_1(1-\delta_2(1-\delta_1(1-\delta_2))))$
...	...	...
∞	$(1-\delta_2)/(1-\delta_1\delta_2)$	$\delta_2(1-\delta_1)/(1-\delta_1\delta_2)$

Table 2: Subgame-perfect partitioning of the surplus as a function of the maximum number of stages n of the alternating-offers game. When the game becomes very long, i.e., when $n \to \infty$, the partitioning of the surplus converges to the partitioning derived by Rubinstein for the infinite-horizon game [17].

Model with a risk of breakdown (Section 7) For risk neutral agents the expected utility of agent i at time t is equal to $x_i p^t$ in this model. If we compare this with the payoff agent i receives in the model with time preferences ($x_i \delta_i^t$, see above), it becomes straightforward to identify the unique SPE strategies for both agents. In particular, the SPE partitioning is given by Table 2 if we replace δ_1 and δ_2 by the continuation probability p. According to Table 2, agent 1 receives $\frac{1}{1+p}$ (and agent 2 the remainder) if n becomes very large. This partitioning is in agreement with predictions for the infinite-horizon game with breakdown [11, pp. 74–77]. Finally, note that the agents agree on an equal split if $p \to 1$ in the infinite-horizon model.

Model with risk averse agents (Section 8) In this case, agent i's payoff at time t is equal to $u_i(x_i)p^t$, where u_i is a concave utility function modeling agent i's attitudes towards risk. In order to calculate the agents' SPE strategies, we repeat the earlier analysis based upon backward induction. The final result is that the SPE partitioning (x_1^*, x_2^*) is given by Table 2 if we replace δ_1 with $u_1^{-1}(p)$ and δ_2 with $u_2^{-1}(p)$.[15] Agent 1 therefore receives a fraction of $\frac{1-u_2^{-1}(p)}{1-u_1^{-1}(p)u_2^{-1}(p)}$ of the surplus (and agent 2 the remainder) if n becomes very large (in agreement with [11, pp. 77–80]). The corresponding partitioning in terms of utility is given by the pair $(u_1(x_1^*), u_2(x_2^*))$.

2. The Evolutionary Algorithm

The pseudo-code of the evolutionary algorithm is given in Table 3. The computer program is written in the Java software language (version 1.2.2). Parameter settings for this algorithm are taken from Table 1. Variants of this algorithm [e.g., (μ, λ) selection instead of $(\mu + \lambda)$ selection] can be implemented easily.

[15]The inverse function of u_i is denoted as u_i^{-1}.

```
1   begin program MAIN
2   generation = 0
3   Generate two populations (pops.) of μ = 100 agents
      parents^i ≡ list of agents in pop. i ∈ {1,2}
      Initialize the chromosome of each agent in parents^i for i = 1,2
4   Calculate fitness parents
      for i = 1,2 do calculateFitness(parents^i)
5   Report results
6   Start main iteration loop
7     generation := generation + 1
8     Generate offspring
        offspring^i ≡ list of offspring for pop. i ∈ {1,2}
        for i = 1,2 do generateOffspring(parents^i)
9     Calculate fitness offspring
        for i = 1,2 do calculateFitness(offspring^i)
10    Collect survivors (parents for the next generation)
        for i = 1,2 do parents^i := selSurvivors(parents^i,offspring^i)
11    Recalculate fitness parents (context has changed)
        for i = 1,2 do calculateFitness(parents^i)
12    Report results
13  Repeat 7 through 12 until the maximum number of generations is reached
14  end program MAIN

    procedure calculateFitness(agents)
1     Select an agent from agents
2     Select opponents (from the other pop.)
        if agent ∈ {parents^1,offspring^1} context := parents^2,
          else context := parents^1
        Select subset of 25 opponents from context
3     Play bargaining game against these opponents
4     Fitness agent is mean utility obtained in these 25 games
5   Repeat 1-4 for all agents in agents

    procedure generateOffspring(parents^i)
1     Select parent from parents^i
2     Form offspring by mutating this parent
3   Repeat 1 and 2 until λ = 100 offspring have been formed
4   Gather all offspring in list offspring^i

    procedure selSurvivors(parents,offspring)
1   Return μ fittest agents from union of parents and offspring
```

Table 3: Pseudo-code for the evolutionary algorithm. Model settings are the same as in Table 1. Names for populations of agents are indicated in italics.

References

1. Th. Bäck. *Evolutionary Algorithms in Theory and Practice.* Oxford University Press, New York, Oxford, 1996.
2. K. Binmore. *Fun and Games.* D.C. Heath and Company, Lexington, MA, 1992.
3. K. Binmore, M. Piccione, and L. Samuelson. Evolutionary stability in alternating-offers bargaining games. *Journal of Economic Theory*, 80:257–291, 1999.
4. K. Binmore and N. Vulkan. Applying game theory to automated negotiation. *Netnomics*, 1(1):1–9, 1999.
5. F.M.T. Brazier, F. Cornelissen, R. Gustavsson, C.M. Jonker, O. Lindeberg, B. Polak, and J. Treur. Negotiating for load balancing of electricity use. In M.P. Papazoglou, editor, *Proceedings of the 18th International Conference on Distributed Computing Systems (ICDCS).* IEEE Computer Society Press, 1998.
6. H. Dawid. *Adaptive Learning by Genetic Algorithms: Analytical Results and Applications to Economic Models.* Lecture Notes in Economics and Mathematical Systems, No. 441. Springer-Verlag, Berlin, 1996.
7. J. Gale, K. Binmore, and L. Samuelson. Learning to be imperfect: The ultimatum game. *Games and Economic Behavior*, 8:56–90, 1995.
8. E.H. Gerding, D.D.B. van Bragt, and J.A. La Poutré. Multi-issue negotiation processes by evolutionary simulation: Validation and social extensions. In *Proceedings of the Workshop on Complex Behaviour in Economics*, Aix-en-Provence, France, May 2000.
9. N. Matos, C. Sierra, and N.R. Jennings. Determining successful negotiation strategies: An evolutionary approach. In *Proceedings of the 3rd International Conference on Multi-Agent Systems (ICMAS-98)*, pages 182–189, Paris, France, 1998.
10. M. Mitchell. *An Introduction to Genetic Algorithms.* The MIT Press, Cambridge, MA, 1996.
11. A. Muthoo. *Bargaining and Applications.* Cambridge University Press, Cambridge, MA, 1999.
12. J.R. Oliver. A machine learning approach to automated negotiation and prospects for electronic commerce. *Journal of Management Information Systems*, 13(3):83–112, 1996.
13. M.J. Osborne and A. Rubinstein. *Bargaining and Markets.* Academic Press, San Diego, CA, 1990.

14. T. Reichmann. Learning and behavioral stability: An economic interpretation of genetic algorithms. *Journal of Evolutionary Economics*, 9(2):225–242, 1999.
15. J.S. Rosenschein and G. Zlotkin. *Rules of Encounter.* The MIT Press, Cambridge, MA, 1994.
16. A.E. Roth, J.K. Murnighan, and F. Schoumaker. The deadline effect in bargaining: Some experimental evidence. *American Economic Review*, 78(4):806–823, 1988.
17. A. Rubinstein. Perfect equilibrium in a bargaining model. *Econometrica*, 50(1):155–162, 1982.
18. R. Selten. Spieltheoretische Behandlung eines Oligopolmodels mit Nachfrageträgheit. *Zeitschrift für die Gesamte Staatswissenschaft*, 12:301–324 and 667–689, 1965.
19. R. Selten. Re-examination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, 4:25–55, 1975.
20. I. Ståhl. *Bargaining Theory.* Stockholm School of Economics, Stockholm, 1972.
21. D.D.B. van Bragt, C.H.M. van Kemenade, and J.A. La Poutré. The influence of evolutionary selection schemes on the iterated prisoner's dilemma. *Computational Economics.* To appear.
22. E. van Damme. *Stability and Perfection of Nash Equilibria.* Springer-Verlag, Berlin, 1991.
23. F. Verhulst. *Nonlinear Differential Equations and Dynamical Systems.* Springer-Verlag, Berlin, 1996.
24. G. Weiss, editor. *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence.* The MIT Press, Cambridge, MA, 1999.