http://www.e-jemed.org/

ISSN : 1298-0137

e - JEMED
The Electronic Journal of Evolutionary Modeling and Economic Dynamics

Article number: 1027

Please cite this article as following:

D.D.B. van Bragt, E.H. Gerding, and J.A La Poutré, 2002, Equilibrium Selection in Alternating-Offers Bargaining Models – The Evolutionary Computing Approach, The Electronic Journal of Evolutionary Modeling and Economic Dynamics, n° 1027, http: //www.e-jemed.org/1027/index.php

---

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

D.D.B. van Bragt
CWI, Amsterdam

E.H. Gerding
CWI, Amsterdam

J.A. La Poutré
CWI, T.U. Eindhoven

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. In our model, each bargaining agent maintains a collection of strategies which is optimized by an evolutionary algorithm (EA). Such evolutionary agents learn in different ways: by selection and reproduction of successful strategies, and by random experimentation (by "mutating" existing strategies) or by recombining or "crossing over" previously-tested strategies.

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 highly nonlinear transients in the ultimatum game, are however observed in several experiments.

Keywords: evolutionary algorithms, evolution strategies, evolutionary game theory, agent-based computational economics, bounded rationality, bargaining, alternating-offers protocol, subgame perfect equilibrium, nonlinear phenomena

JEL: C61, C63, D49

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

D.D.B. van Bragt* E.H. Gerding* J.A. La Poutré*,†

*CWI, Centre for Mathematics and Computer Science

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

†School of Technology Management, Eindhoven University of Technology

P.O. Box 513, 5600 MB Eindhoven, The Netherlands.

E-mail: {bragt, egerding, hlp}@cwi.nl.

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. In our model, each bargaining agent maintains a collection of strategies which is optimized by an evolutionary algorithm (EA). Such evolutionary agents learn in different ways: by selection and reproduction of successful strategies, and by random experimentation (by "mutating" existing strategies) or by recombining or "crossing over" previously-tested strategies.

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 highly nonlinear transients in the ultimatum game, are however observed in several experiments.

JEL Classification: C61, C63, D49

Keywords: evolutionary algorithms, evolution strategies, evolutionary game theory, agent-based computational economics, bounded rationality, bargaining, alternating-offers protocol, subgame perfect equilibrium, nonlinear phenomena

1 Introduction

Recently, interest in the development of trading and negotiating agents has surged among economists and computer scientists (Binmore and Vulkan, 1999). A nice example of the potential of automated negotiation is given by Brazier et al. (1998). They describe 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. Other practical applications of distributed multi-agent systems are surveyed by Weiss (1999, Ch. 9).

The rapid establishment of a global communication network (in the form of the Internet) together with the development of standard negotiation protocols (Rosenschein and Zlotkin, 1994) 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 model an adaptive agent as a collection of strategies which is optimized by an evolutionary algorithm (EA) (Mitchell, 1996; Bäck, 1996). EAs transfer the principles of natural evolution, first discovered by Darwin, to a computational setting. These algorithms have been used in the past, with considerable success, to solve difficult optimization problems. Examples include problems with huge search spaces, multiple local optima, discontinuities, and noise (Mitchell, 1996; Bäck, 1996). Adaptive agents learn in different ways in an evolutionary setting: by selection and reproduction of successful strategies, and by random experimentation (by "mutating" existing strategies) or by recombining or "crossing over" previously-tested strategies.

An evolutionary agent, as described above, is boundedly rational for several reasons. Firstly, such an agent does not base its decisions on a formal analysis of the game, but, instead, learns by trial-and error. Secondly, the opponent (or opponents) of the evolutionary agent are not modelled explicitly. Thirdly, the only feedback that is used by the evolutionary agent is the performance (payoff) of its strategies. Using this feedback, the strategies with a low payoff are replaced by new strategies in the course of time. Fourthly, an evolutionary agent only maintains a limited collection of game strategies, i.e., not all possible strategies are evaluated.

Previous research has demonstrated that, despite these limitations, evolutionary agents can develop highly effective negotiation strategies. An early example was given by Oliver (1996). He performed computer simulations of both distributive (i.e., single-issue) and integrative (i.e., multiple-issue) "alternating-offers" negotiations. In Oliver's model, the bargaining strategies are represented as binary-coded strings. 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 are then updated in successive generations by a genetic algorithm (GA).

More elaborate strategy representations were proposed and evaluated by Matos et al. (1998). Offers and counter offers are generated in their model by a linear combination of simple bargaining tactics (time-dependent, resource-dependent, or behavior-dependent tactics). As in (Oliver, 1996), 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 strategies, which were simultaneously evolved by a GA.

We intend to bridge the gap between the above-described computer experiments and the analysis of bargaining by game theorists (Ståhl, 1972; Rubinstein, 1982; Osborne and Rubinstein, 1990). 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 random strategies. As a consequence, many different paths through the game tree will be explored (i.e., many subgames will be sampled). Only those strategies which are relatively successful in many different subgames will be selected as parents for the next generation of strategies. 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) (Selten, 1965, 1975). 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 (1982) successfully applied this notion of subgame-perfection to bargaining games. 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]

Our computational experiments indicate that evolutionary agents, with a bounded rationality, may actually display subgame-perfect behavior in the alternating-offers game. Moreover, we encounter phenomena beyond the reach

[1] That is, games with a tree structure (Osborne and Rubinstein, 1990).

[2] The finite-horizon variant of Rubinstein's game (which we use in our computer simulations) has been analyzed earlier by Ståhl (1972).

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 always 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 a description of the bargaining model that we investigate in this paper. An overview of the setup of the computational experiments is then given in Section 3. Sections 4 contains an analysis of the computational results. Section 5 gives some pointers to related work and Section 6 concludes.

2 The Alternating-Offers Bargaining Model

We use a finite-horizon variant of Rubinstein's bargaining model (1982). In this game, two agents bargain with each other over the partitioning of a constant surplus.[3] Offers are made at discrete points in time: namely, at times $t = 0, \ldots, (n-1)$, where n is the maximum number of stages of the bargaining game. We denote an offer made at time t as $o(t)$. An offer $o(t)$ specifies the share of the surplus that the initial proposer ("agent 1") receives if the offer is accepted at time t (agent 2 then receives $1 - o(t)$).

The two agents bargain in an alternating fashion. At $t = 0$, agent 1 makes the first 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. We set $n = 10$ in our computer experiments (unless stated otherwise).

Following Rubinstein (1982), 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 received by agent i. Agents thus experience time pressure because they prefer to reach an agreement early.

Subgame-perfect equilibrium strategies for this finite-horizon game can be derived by using a backward-induction approach (van Damme, 1991, Ch. 1). The SPE partitioning (x_1^*, x_2^*) as a function of the game length is listed in Table 1.[4]

[3] Without loss of generality, we set the size of this surplus equal to unity.

[4] A formal derivation of these expressions can be found in (van Bragt et al., 2000, Appendix 1).

In equilibrium, agent 1 demands a share of $x_1^*(n)$ in the first round and agent 2 immediately accepts this proposal [receiving $x_2^*(n) = 1 - x_1^*(n)$]. To be expected, the partitioning of the surplus converges to the partitioning derived by Rubinstein (1982) for the infinite-horizon game. 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 for agent 1 (x_1^*)	SPE share for 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 1: Subgame-perfect partitioning of the surplus as a function of the maximum number of stages of the alternating-offers game (n).

To derive the SPE strategies, it is assumed 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 boundedly-rational agents.

3 The Evolutionary Model

We here describe an abstract model of two adaptive bargaining agents who are updating their strategies over time. In our model, each bargaining agent maintains its own collection of strategies. Each collection of bargaining strategies is then evolved over time by an evolutionary algorithm (EA). Section 3.1 gives an outline of the EA and discusses how our evolutionary system can be interpreted as a model for economic learning processes. 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 3.3-3.5.

3.1 The Evolutionary Algorithm

We model an adaptive agent as a collection of strategies which is optimized by an evolutionary algorithm (EA) (Mitchell, 1996; Bäck, 1996). EAs transfer the principles of natural evolution, first discovered by Darwin, to a computational setting. These algorithms have been used in the past, with considerable success, to solve difficult optimization problems. Examples include problems with huge search spaces, multiple local optima, discontinuities, and noise (Mitchell, 1996; Bäck, 1996).

As in natural ecosystems, EAs typically evolve a population of individuals. Here, each individual is a bargaining strategy of the adaptive agent (see Section 3.2). Our evolutionary model consists of two co-evolving agents (where each agent maintains its own collection of strategies). We assume that one of the agents, denoted as “agent 1”, has the privilege 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.

Like in nature, the survival probability of each bargaining strategy depends on its fitness (the “survival of the fittest” concept). During the fitness evaluation, each strategy competes against a group of opponent strategies who are drawn at random (without replacement) from the population of strategies of the other agent.[5] The strategy’s fitness is then equal to the mean payoff obtained against these opponent strategies.

Using this fitness information, 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 strategies is determined by competition between the strategies in the two populations. In the next stage (see Fig. 1), “offspring” strategies are created. An offspring strategy is generated in two steps. First, a strategy in the parental population is (randomly, with replacement) selected. This strategy is then mutated to create a new offspring strategy (the mutation model is specified in detail in Section 3.4). The fitness of the new offspring is evaluated by interaction with the parental strategies. An economic interpretation of this parent-offspring interaction is that new strategies need to be able to compete with existing or “proven” strategies before they gain access to the agent’s strategy pool.[6] In the final stage of the iteration (see Fig. 1), the

[5]Obviously, the size of the group of opponent strategies is an important model parameter. We sample 25 (out of a total of 100) opponent strategies (see Table 2). If the size of the opponent group becomes much smaller (e.g., equal to unity), the fitness determination becomes very noisy.

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

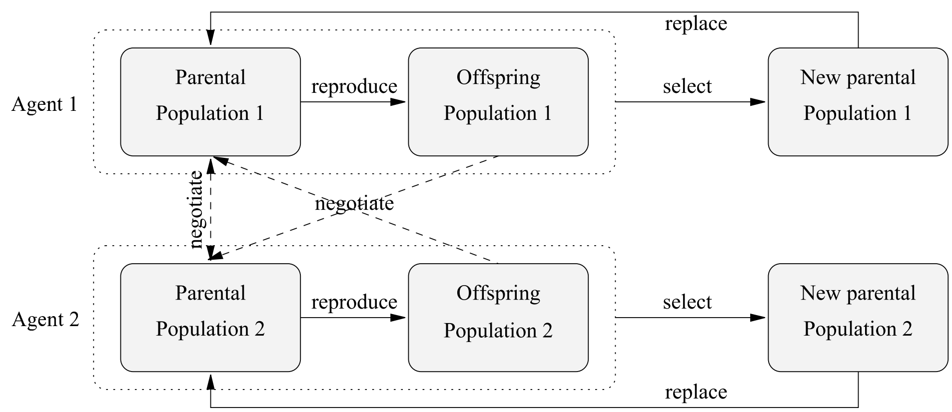

Figure 1: Iteration loop of the evolutionary algorithm (EA). This algorithm updates the populations of strategies which are used by the two adaptive agents.

fittest strategies 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 2. Pseudocode of the computational model can be found in Appendix A.

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 2: Settings of the evolutionary model.

3.2 Genetic Representation

In our model, each strategy specifies a list of offers and thresholds for the different negotiation rounds. The thresholds determine whether an offer of the other party is accepted or rejected: If the proposed fraction of the surplus falls below the threshold the offer is refused (and a counter offer is made in the next round); otherwise an agreement is reached.

Each strategy is encoded as a sequence of real-coded genes (together called a "chromosome") in our evolutionary system. This representation is depicted schematically in Fig. 2. Notice that in each round, the strategy specifies either an offer or a threshold, depending on whether the agent who uses the strategy proposes or receives an offer in that round. The length l of each chromosome is thus

Strategy agent 1:	$o(t=0)$	$\tau(t=1)$	$o(t=2)$	$\tau(t=3)$	$\ldots$
Strategy agent 2:	$\tau(t=0)$	$o(t=1)$	$\tau(t=2)$	$o(t=3)$	$\ldots$

Figure 2: The strategies for agent $i \in \{1,2\}$ specify a sequence of offers $o(t)$ and thresholds $\tau(t)$ for rounds $t \in \{0,1,2,\ldots,n-1\}$ of the negotiation.

equal to the number of rounds (n). Because the size of the bargaining surplus is equal to 1, the offers and thresholds are also restricted to the unity interval. The agents' strategies are initialized at the beginning of each experiment by drawing a random number in the unit interval for each gene (from a uniform distribution).

The above-described representation of the bargaining strategies, which was originally proposed by Oliver (1996), is rather elementary. In fact, the offers made in the consecutive rounds are fully pre-programmed in the genes. The evolution of more reactive bargaining strategies, i.e., strategies which are able to display behavior which is conditional on the opponent's moves, is studied in two companion papers (van Bragt and La Poutré, 2002a,b). Readers who are interested in the development of more complex and powerful strategy representations are referred to these works for further details.

3.3 Selection Scheme

Selection is performed using the $(\mu+\lambda)$-ES selection scheme (Bäck, 1996). In conventional notation, μ is the number of parents and λ is the number of generated offspring ($\mu = \lambda = 100$, see Table 2). The μ survivors with the highest fitness are selected from the union of parental and offspring strategies. This selection scheme is therefore an example of an "overlapping generations" model, in which successful strategies can survive for multiple generations.[7]

An offspring strategy is generated in two steps. First, a strategy in the population is (at random, with replacement) selected to be a parent. The chromosome of this parental strategy is then mutated to generate a new offspring strategy (the mutation model is specified below in Section 3.4). We set the parent-to-offspring

[7]A nonoverlapping generations model, in which all parents are discarded after one generation, is investigated by van Bragt et al. (2000). A probabilistic variant of $(\mu+\lambda)$-ES selection is also studied in this reference.

ratio equal to unity (i.e., $\mu = \lambda$).[8]

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. It is important to note that, in our model, the agent's strategies are not binary strings (as in most GA implementations) but, instead, consist of strings of real-coded numbers. A subfield of evolutionary computation, called "evolution strategies" (ES), has developed the proper evolutionary techniques to adapt such real-coded strings (see Bäck (1996) for an overview). The standard approach in the field of ES is to mutate each gene (consisting of a real-coded number) by adding a zero-mean Gaussian variable to the gene's value. This approach is also used in our evolutionary model (with positive results).

More formally, we create the offspring's genes x_i by adding a zero-mean Gaussian variable, with standard deviation $\sigma_i = 0.1$ [i.e., $N_i(0, 0.1)$],[9] to each corresponding gene x_i of the parent.[10] All offspring genes with a value larger than unity (or smaller than zero) are reset to unity (respectively zero).[11]

3.5 Recombination Model

The recombination (or "crossover") operator exchanges parts of the parental chromosomes to produce new offspring. This facilitates a rapid exchange of genetic information in an agent's strategy pool. Recombination of genetic information has proven to be a very effective search operator if the individuals are binary-coded (Mitchell, 1996). Following this lead, several recombination models have also been proposed for evolutionary models with real-coded individuals (Bäck, 1996).

[8] Several experiments are reported by van Bragt et al. (2000) in which this ratio is not equal to unity (to determine the influence of the selection intensity on the equilibrium selection process).

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

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

[11] An alternative approach would be to enable individual strategies 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 (Bäck, 1996, pp. 71-73). This model, which is studied further by van Bragt et al. (2000), allows an evolutionary self-adaptation of both the genes and the corresponding standard deviations at the same time.

We performed experiments with two recombination models which are frequently used in the field of ES: discrete recombination and intermediate recombination (Bäck, 1996, 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.

4 Results

In Section 4.1, both agents have identical discount factors (i.e., $\delta_1 = \delta_2 \equiv \delta$). Results for $\delta_1 \neq \delta_2$ are presented in Section 4.2. During the evolutionary experiments, we monitor the performance of the two adaptive agents. We define the performance (fitness) of an adaptive agent as the mean fitness across all (100) strategies maintained by the agent.

4.1 Symmetric Time Preferences

We first investigate two extreme cases. In Section 4.1.1 we set $\delta = 0$. Agents then receive nothing if they do not reach agreement in the very first round. In payoff terms, this situation is equivalent with the ultimatum game. Another extreme case is obtained by setting $\delta = 1$. In this case, analyzed in Section 4.1.2, the agents are payoff-indifferent between reaching a deal sooner or later (provided $t < n$). Results for intermediate values of δ ($0 < \delta < 1$) are summarized in Section 4.1.3.

4.1.1 $\delta = 0$ (The Ultimatum Game)

Figure 3 shows the evolution of the agents' fitnesses in the ultimatum game (for a typical experiment). Game theory predicts that the proposer (i.e., agent 1) demands the whole surplus, which the responder (i.e., agent 2) accepts.[12] This unique (subgame-perfect) equilibrium indeed appears to be an attractor for the evolutionary system: the fitness of agent 1 increases rapidly initially, whereas the fitness of agent 2 is decreasing at the same time.

Figure 3 also reveals, however, that there is no stable convergence to subgame-perfect behavior. Instead, highly nonlinear transients are visible in Fig. 3. These transients start directly after the SPE partitioning is reached by the adaptive agents. At this point, first reached after ≈ 175 generations in Fig. 3, agent 2 becomes (payoff) indifferent between accepting or refusing agent 1's extreme offer (the result is the same, agent 2 receives nothing). The mutation process continues to create offspring strategies with a threshold larger than zero in agent 2's strategy

[12] To see this, let δ_1 and δ_2 approach zero in Table 1.

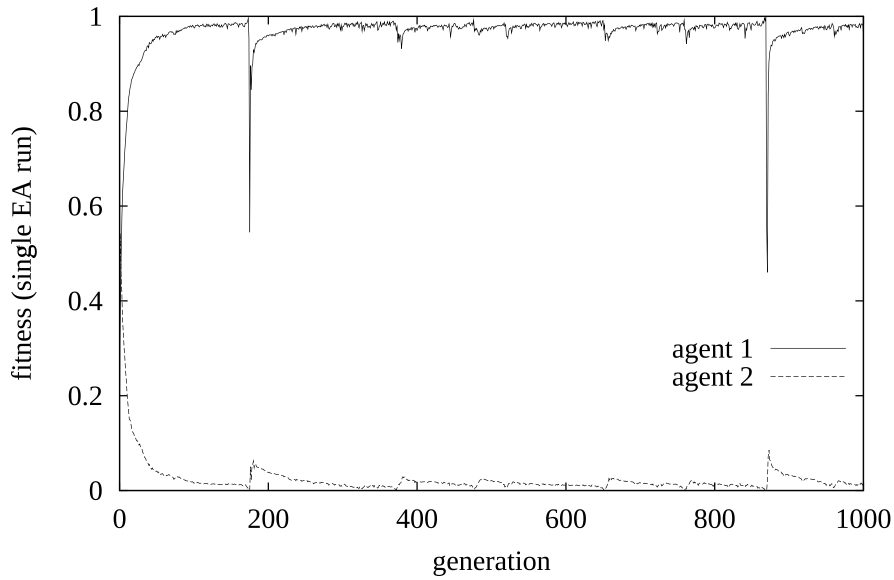

Figure 3: Evolution of the agents' fitnesses in the ultimatum game (a typical EA run is shown).

pool in this case. These strategies have the same fitness as their all-accepting counterparts. Therefore, some of them invade the strategy pool of agent 2. This results in a significant number of disagreements and a sharp drop in fitness for agent 1. Consequently, some strategies in agent 1's strategy pool decrease their offer in order to stop this process and the fitness of agent 2 increases slightly. The race between agent 1 and agent 2 then starts all over again, and the process repeats itself (see Fig. 3).[13]

The influence of changes in the agents' EA on the equilibrium-selection process is investigated in detail in (van Bragt et al., 2000, Ch. 5). An important conclusion of this study is that the specific EA used by the adaptive agents can have a strong impact on the (long-term) partitioning of the bargaining surplus in ultimatum game situations. For example, if the agents use an EA with "non-overlapping generations" (so that all strategies from the previous generation are discarded) convergence to equilibria which are not subgame-perfect can occur in the long run.[14] The same may also happen when the agents use an EA with stochastic se-

[13]Nonlinear population dynamics has also been encountered in co-evolving populations of predators and preys in natural ecosystems. A simple mathematical model describing such a system has been proposed by Lotka and Volterra.

[14]Convergence to an equilibrium which is not subgame-perfect in the ultimatum game has been reported before in the field of evolutionary game theory (Gale et al., 1995). In (Gale et al., 1995)

lection or an EA with a small selection pressure. We refer the interested reader to (van Bragt et al., 2000, Ch. 5) for more details.

4.1.2 $\delta = 1$ (Time Indifference)

Figure 4 shows the evolution of the fitness of agent 1 (averaged over 25 EA runs) in the *n*-stage alternating-offers game (without payoff discounting). Game theory

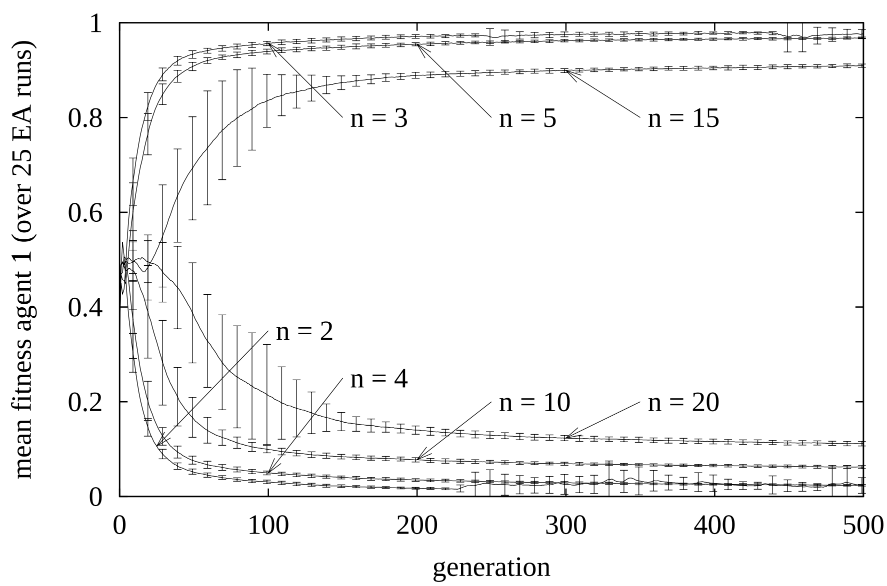

Figure 4: Evolution of the fitness of agent 1 in *n*-stage alternating-offers games without payoff discounting. In the long run, agent 1 receives the largest share of the surplus if he has the opportunity to make the last offer (i.e., when *n* is odd). Exactly the opposite happens when agent 2 is last in turn (i.e., when *n* is even).

predicts that the last agent in turn receives the entire bargaining surplus in this case.[15] Hence, we would expect that agent 1 receives the entire surplus if *n* is odd and nothing if *n* is even. This tendency is indeed clearly visible in Fig. 4, even for games as long as 20 rounds.

The timing of the agreements is not uniquely defined at subgame perfect equilibrium in the absence of time pressure.[16] It is therefore of interest to investigate

the evolution of strategies for the ultimatum game is governed by a nonoverlapping generations variant of the replicator dynamics (Binmore, 1992, Ch. 9).

[15]To see this, let δ_1 and δ_2 approach unity in Table 1.

[16]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

the timing of the agreements in the evolutionary system. We observe that in the evolutionary 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 (Roth et al., 1988).

In the very long run, the SPE is sometimes reached by the adaptive agents. The last agent in turn receives the entire surplus in this case, whereas his opponent receives nothing. Strategies with a non-zero threshold for $t = n - 1$ then invade the strategy pool of the agent who receives the final "take-it-or-leave-it" offer (see Section 4.1.1). The last agent in turn then avoids the occurrence of a large number of disagreements by rapidly decreasing its offers and thresholds in earlier rounds. As a consequence, more agreements will temporarily occur in earlier rounds after the SPE has been reached. Afterwards, the last agent in turn will again delay the agreements until the very last round, etc.

4.1.3 $0 < \delta < 1$

We study the partitioning of the surplus for a wider range of discount factors in Figs. 5 and 6. The agents' fitnesses are measured after 500 generations (and averaged over 25 EA runs) in these figures. In the long run, agent 1 often receives more than game theory predicts, whereas agent 2 negotiates relatively poor deals. This effect is particularly clear in case of strong time pressure (i.e., a small δ). We observe in the computer experiments that almost all agreements are reached immediately in this case (e.g., 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). The short duration of the game is exploited effectively by agent 1: like in the ultimatum game, this agent demands (and receives) a large share of the surplus.

Figures 5 and 6 also 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 6 shows for instance that agent 2 does not fully exploit his last-mover advantage under these circumstances (his mean fitness is far below the SPE level). This effect can be explained by the boundedly-rational behavior of

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.

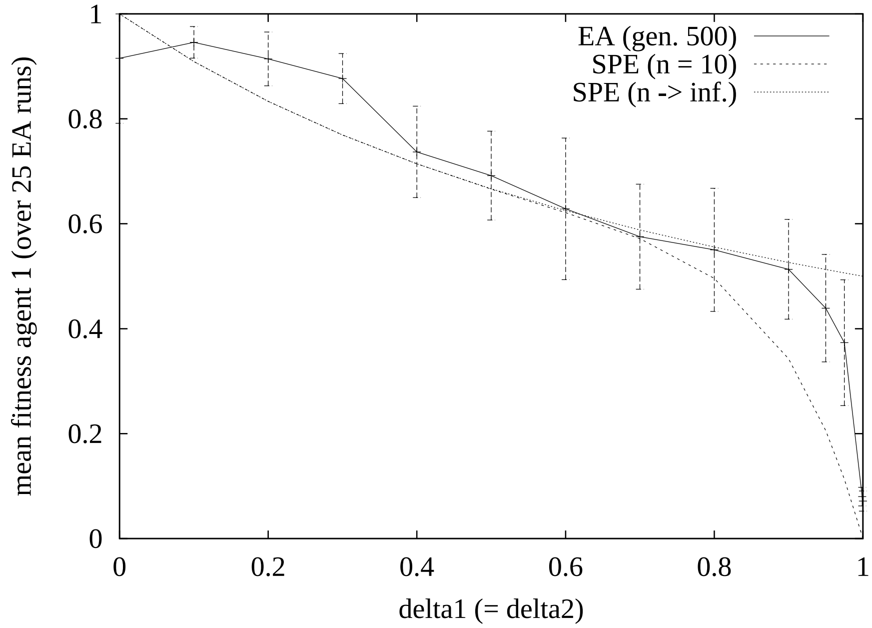

Figure 5: Performance of agent 1 as a function of the discount factor. Game theoretic predictions for the 10-stage game and the infinite-horizon game are also shown for comparison.

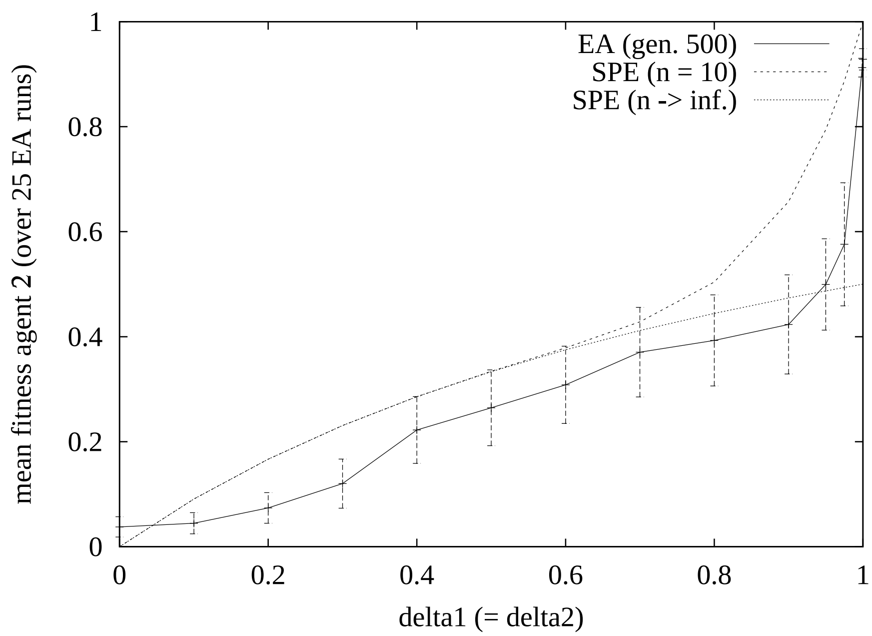

Figure 6: Performance of agent 2 as a function of the discount factor.

the adaptive agents. These agents do not reason backwards from the deadline, 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 modelling assumption if the agents do not perceive the finite deadline of the game. Figures 5 and 6 indeed show that the experimental outcome is predicted quite well (for δ up to 0.9) by theoretical predictions for an infinite-horizon game.

4.2 Asymmetric Time Preferences

Figure 7 shows the long-run performance of the adaptive agents in case of asymmetric time preferences. Agent 1's discount factor (δ_1) is set equal to 0.6 whereas

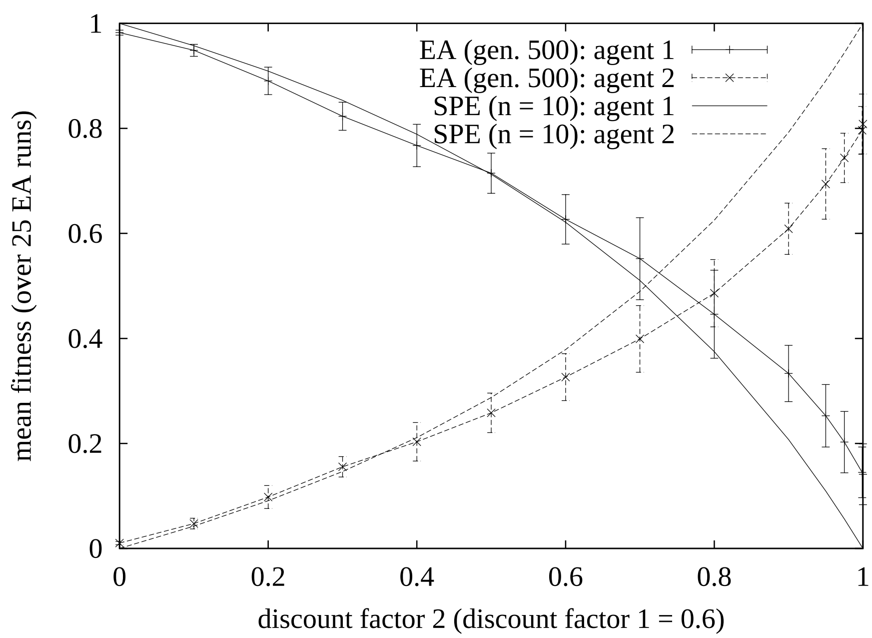

Figure 7: Long-term fitnesses of agent 1 and agent 2 in case of asymmetric time preferences ($\delta_1 = 0.6$ and δ_2 is varied between zero and unity).

the discount factor of agent 2 (δ_2) is varied between zero and unity. The fitnesses of the agents converge within 50-150 generations to the values reported in Fig. 7. Note that the performance of agent 2 is not as good as predicted by game theory when $\delta_2 > \delta_1$, while agent 1 actually does better. This effect becomes especially clear if $\delta_2 = 1$. We will study this case in more detail below.

When $\delta_2 = 1$, agent 2 experiences no explicit time pressure to reach an early agreement. Time pressure is, on the other hand, relatively large for agent 1 (his payoff diminishes proportional to 0.6^t as a function of the round number t). This reduces the evolutionary pressure against strategies in agent 1's strategy pool with a large threshold or offer gene (for large t). In the experiments we even observe that these genes evolve to random values (in the unit interval) for $t \geq 4$.

Now assume that agent 2 tries to exploit his bargaining power by delaying agreements. Agent 2 will then encounter an opponent who is using a random strategy in later rounds. This deprives agent 2 partly of his bargaining power: agent 2 cannot force his indifferent opponent to adjust his behavior in later rounds. In fact, exactly the opposite occurs in the evolutionary system. In an attempt to avoid the occurrence of disagreements, agent 2 reduces his acceptance threshold and increases his offer (to agent 1) in later rounds.[17]

Experiments with alternative evolutionary models (see van Bragt et al. (2000)) lead to similar results for $\delta_2 > \delta_1$. Hence, the deviations from game-theoretic predictions in the computational experiments cannot be attributed to the specific settings of the EA (which was used to generate Fig. 7).

5 Related Work

Several authors have adopted and further extended our evolutionary framework.[18] The extension to (much more complex) multi-issue negotiations is presented in (Gerding et al., 2000). In multi-issue negotiations not just one issue (like the price of a product) is important, but other aspects are also taken into account (for instance accessories, quality, delivery time, etc.). A key advantage of these multi-issue negotiations is that often mutually beneficial outcomes can be obtained if both parties concede on the proper issues. The complexity of the bargaining problem increases rapidly, however, if the number of issues becomes large.

Gerding et al. (2000) show, however, that adaptive agents (using EAs) are able to generate Pareto-efficient outcomes for bargaining problems with up to 8 issues. The decision-making process of the adaptive agents is also extended by Gerding et al. (2000) by allowing the agents to use a "fairness" norm in the negotiations. This concept plays an important role in real-life negotiations and experimental economics.

[17] $\tau(t = 0)$ evolves for instance to 0.78 ± 0.14 in the strategy pool of agent 2, whereas $\tau(t = 4) = 0.64 \pm 0.26$ and $\tau(t = 8) = 0.32 \pm 0.25$ in the long run. Agent 2's offer genes evolve to values of $o(t = 1) = 0.10 \pm 0.09$, $o(t = 5) = 0.24 \pm 0.32$, and $o(t = 9) = 0.45 \pm 0.31$.

[18] Using the earlier published technical report (van Bragt et al., 2000) as their reference.

Gerding and La Poutré (2002) study an ultimatum game in which the bargaining agents have multiple bargaining opportunities. Such a game can be considered as an abstract model of a competitive market, where a buyer can for instance try different sellers before making a purchase. Gerding and La Poutré (2002) present results of various evolutionary simulations for this market game.

In a recent series of papers, Nawa, Shimohara, and Katai also study several variants of our model. A bargaining model with 3 adaptive agents is investigated in (Nawa et al., 2001a). A multi-issue bargaining model in which the issues are negotiated sequentially is studied in (Nawa et al., 2001b) (the issues are negotiated simultaneously in Gerding et al. (2000)). Nawa et al. (2001c) demonstrate that fair agreements can evolve if there exists uncertainty about which agent starts the bargaining game. Finally, the effect of evolutionary learning parameters on the bargaining outcome is assessed in (Nawa et al., 2001d).

The above-described works all use the strategy representation proposed by Oliver (1996). This representation is quite static, however, since the offers made in the consecutive rounds are fully pre-programmed in the genes. The evolution of more reactive bargaining strategies, i.e., strategies which are able to display behavior which is conditional on the opponent's moves, is studied by van Bragt and La Poutré (2002a,b). The bargaining strategies are represented in these works by a special kind of finite automata. Computational experiments show that adaptive agents (based upon these automata) are able to discriminate successfully between different (static or co-evolving) opponents, although they receive no explicit information about the strategy, identity or preferences of their opponents. Obviously, these results are important for the further development of adaptive agents for real-life applications.

Carmel and Markovitch (1996, 1999) have proposed an interesting model-based approach for learning effective strategies in multi-agent systems. They restrict the agents' strategies to deterministic finite automata and show that a best response strategy for a given opponent (with a known strategy) can be derived efficiently. They furthermore present an unsupervised learning algorithm that infers a model of the opponent's automaton from its input/output behavior. These techniques were applied successfully to the iterated prisoner's dilemma game. Although the framework of Carmel and Markovitch appears to be promising in case of fixed opponents, the case of non-stationary opponents is not covered yet by their methods. A second limitation of their approach is that explicit information about the identity of the opponents should be available in a setting with multiple opponents (to facilitate the development of separate finite automaton models for the different opponents). Information about the identity (or preferences) of the opponents is not used (or needed) in our evolutionary model. For example, the only feedback used in the multi-opponent model studied by van Bragt and La Poutré (2002a,b) is the average score obtained by the automata against the different opponents.

6 Conclusions and Future Work

We have studied the dynamic and equilibrium-selecting behavior of a multi-agent system consisting of adaptive bargaining agents. In our model, each bargaining agent maintains a collection of strategies which is optimized by an evolutionary algorithm (EA). Such evolutionary agents learn in different ways in an evolutionary setting: by selection and reproduction of successful strategies, and by random experimentation (by "mutating" existing strategies) or by recombining or "crossing over" previously-tested strategies. Negotiations between the adaptive agents are governed by a finite-horizon version of Rubinstein's well-known "alternating-offers" protocol.

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 strong or weak) significant deviations from game-theoretic predictions can occur, however.

A good example is the case of extreme time pressure. In this case, highly nonlinear transients can occur if the deal reached by the adaptive agents approaches the extreme outcome predicted by game theory. Two other experimental observations should also 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 when the agents do not perceive the deadline of the negotiations. Second, we observe (and explain) discrepancies between theory and experiment if the agents experience an unequal time pressure.

More in general, this work presents a systematic validation of evolutionary and computational techniques in the field of bargaining. Our model has also served as a starting point for further explorations (see (Gerding et al., 2000; Nawa et al., 2001a,b,c,d; Gerding and La Poutré, 2002; van Bragt and La Poutré, 2002a,b)). Several important topics have been addressed in these works: complex multi-issue and multi-opponent bargaining problems, economic modelling issues, learning by co-evolution, the development of powerful bargaining strategies, etc. We hope that these different lines of research will be extended further in future works.

An interesting topic for further studies is the impact of asymmetric speeds of learning on the (long-run) behavior of adaptive agents. In the model presented in this paper, the mutation and selection processes are the same for both agents. It would thus be interesting to study what happens when, for one of the agents, there is a kind of inertia, so that, for example, the offspring strategies are closer to the parental strategies for one agent than for the other one. This may induce a kind of (asymmetric) delay in the adaptation process, which would be an interesting topic for further studies.

Acknowledgements

This paper has been presented at the 6^{th} International Conference of the Society for Computational Economics on Computing in Economics and Finance (CEF'2000) (Barcelona, Catalonia, Spain, July 6-8, 2000) and the 1^{st} World Congress of the Game Theory Society (GAMES'2000) (Bilbao, Spain, July 24-28, 2000). The authors would like to thank all conference participants for their feedback. Special thanks are due to Stefan Napel for his constructive comments and useful suggestions. The constructive feedback of two anonymous referees is also appreciated.

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.

A Appendix: Pseudo Code

The pseudo-code of the evolutionary model is given in Table 3. The computer program is written in the Java software language (version 1.2.2). Parameter settings are taken from Table 2.

```
1   begin program MAIN
2   generation = 0
3   Generate two populations (pops.) of μ = 100 strategies
      parents^i ≡ list of strategies in pop. i ∈ {1,2}
      Initialize the chromosome of each strategy 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(strategies)
1     Select a strategy from strategies
2     Select opponent strategies (from the other pop.)
        if strategy ∈ {parents^1,offspring^1} context := parents^2,
          else context := parents^1
        Select subset of 25 opponent strategies from context
3     Play bargaining game against these opponents
4     Fitness strategy is mean payoff obtained in these 25 games
5   Repeat 1-4 for all strategies in strategies

    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 strategies from union of parents and offspring
```

Table 3: Pseudo-code for the evolutionary model. Model settings are the same as in Table 2. Names for populations of strategies are indicated in italics.

References

Bäck, T. (1996). *Evolutionary Algorithms in Theory and Practice*. Oxford University Press, New York and Oxford.

Binmore, K. (1992). *Fun and Games*. D.C. Heath and Company, Lexington, MA.

Binmore, K. and Vulkan, N. (1999). Applying game theory to automated negotiation. *Netnomics*, 1(1):1–9.

Brazier, F., Cornelissen, F., Gustavsson, R., Jonker, C., Lindeberg, O., Polak, B., and Treur, J. (1998). Negotiating for load balancing of electricity use. In Papazoglou, M., editor, *Proceedings of the 18th International Conference on Distributed Computing Systems (ICDCS)*. IEEE Computer Society Press.

Carmel, D. and Markovitch, S. (1996). Learning models of intelligent agents. In *Proceedings of the Thirteenth National Conference on Artificial Intelligence*, Portland, Oregon.

Carmel, D. and Markovitch, S. (1999). Exploration strategies for model-based learning in multiagent systems. *Autonomous Agents and Multi-Agent Systems*, 2(2):141–172.

Gale, J., Binmore, K., and Samuelson, L. (1995). Learning to be imperfect: The ultimatum game. *Games and Economic Behavior*, 8:56–90.

Gerding, E. and La Poutré, J. (2002). Multi-issue bargaining with multiple encounters: An evolutionary social simulation. In *Proceedings of the Workshop on Economics with Heterogeneous Interacting Agents (WEHIA-2002)*, Trieste, Italy.

Gerding, E., van Bragt, D., and La Poutré, J. (2000). Multi-issue negotiation processes by evolutionary simulation: Validation and social extensions. Technical Report SEN-R0024, CWI, Amsterdam. Presented at the Workshop on Complex Behavior in Economics, Aix-en-Provence, France, May 4-6, 2000.

Matos, N., Sierra, C., and Jennings, N. (1998). 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.

Mitchell, M. (1996). *An Introduction to Genetic Algorithms*. The MIT Press, Cambridge, MA.

Nawa, N., Shimohara, K., and Katai, O. (2001a). Does diversity lead to morality? On the evolution of strategies in a 3-agent alternating-offers bargaining model. In *Workshop on Evolutionary Computation and Multi-Agent Systems (ECOMAS) at the Genetic and Evolutionary Computation Conference (GECCO-2001)*, pages 317–320, San Francisco, USA.

Nawa, N., Shimohara, K., and Katai, O. (2001b). Equilibrium selection in a sequential multi-issue bargaining model with evolutionary agents. In *Proceedings of the 2nd Asia-Pacific Conference on Intelligent Agent Technology (IAT-2001)*, Maebashi, Japan.

Nawa, N., Shimohara, K., and Katai, O. (2001c). On fairness in an alternating-offers bargaining model with evolutionary agents. In *Proceedings of the 4th Pacific Rim International Workshop on Multi-Agents (PRIMA-2001)*, pages 63–77, Taipei, Taiwan.

Nawa, N., Shimohara, K., and Katai, O. (2001d). On the effect of evolutionary agents learning parameters in an alternating-offers bargaining model. In *Workshop on Economic Agents, Models, and Mechanisms, 17th International Joint Conference on Artificial Intelligence (IJCAI-01)*, Seattle, USA.

Oliver, J. (1996). A machine learning approach to automated negotiation and prospects for electronic commerce. *Journal of Management Information Systems*, 13(3):83–112.

Osborne, M. and Rubinstein, A. (1990). *Bargaining and Markets*. Academic Press, San Diego, CA.

Rosenschein, J. and Zlotkin, G. (1994). *Rules of Encounter*. The MIT Press, Cambridge, MA.

Roth, A., Murnighan, J., and Schoumaker, F. (1988). The deadline effect in bargaining: Some experimental evidence. *American Economic Review*, 78(4):806–823.

Rubinstein, A. (1982). Perfect equilibrium in a bargaining model. *Econometrica*, 50(1):155–162.

Selten, R. (1965). Spieltheoretische Behandlung eines Oligopolmodels mit Nachfrageträgheit. *Zeitschrift für die Gesamte Staatswissenschaft*, 12:301–324 and 667–689.

Selten, R. (1975). Re-examination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, 4:25–55.

Ståhl, I. (1972). *Bargaining Theory*. Stockholm School of Economics, Stockholm.

van Bragt, D., Gerding, E., and La Poutré, J. (2000). Equilibrium selection in alternating-offers bargaining models: The evolutionary computing approach. Technical Report SEN-R0013, CWI, Amsterdam.

van Bragt, D. and La Poutré, J. (2002a). Co-evolving automata negotiate with a variety of opponents. In *Proceedings of the 2002 Congress on Evolutionary Computation (CEC'2002)*, volume 2, pages 1426–1431. IEEE Press.

van Bragt, D. and La Poutré, J. (2002b). Why agents for automated negotiations should be adaptive. Technical Report SEN-R0208, CWI, Amsterdam. Presented at the 7^{th} International Conference of the Society for Computational Economics on Computing in Economics and Finance (CEF'2001) (New Haven, CT, USA, June 28-29, 2001).

van Damme, E. (1991). *Stability and Perfection of Nash Equilibria*. Springer-Verlag, Berlin.

Weiss, G., editor (1999). *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. The MIT Press, Cambridge, MA.