Co-evolutionary Strategies for an Alternating-Offer Bargaining Problem

Nanlin Jin
Department of Computer Science
University of Essex
Colchester CO4 3SQ, UK
njin@essex.ac.uk

Edward Tsang
Department of Computer Science
University of Essex
Colchester CO4 3SQ, UK
edward@essex.ac.uk

Abstract- In this paper, we apply an Evolutionary Algorithm (EA) to solve the Rubinstein's Basic Alternating-Offer Bargaining Problem, and compare our experimental results with its analytic game-theoretic solution. The application of EA employs an alternative set of assumptions on the players' behaviors. Experimental outcomes suggest that the applied co-evolutionary algorithm, one of Evolutionary Algorithms, is able to generate convincing approximations of the theoretic solutions. The major advantages of EA over the game-theoretic analysis are its flexibility and ease of application to variants of Rubinstein Bargaining Problems and complicated bargaining situations for which theoretic solutions are unavailable.

1 Introduction

There are several methodologies applied for studying games. One of them is to evolve players' strategies in a way that simulates the natural evolution. Evolutionary Algorithm (EA) have proven effective for a wide variety of problems.

The purpose of our research is to apply Evolutionary Algorithms, specifically Co-evolutionary Algorithms, to solve an Alternating-Offer Bargaining Problem. To this problem, the assumptions of the game-theoretic solutions Subgame Perfect Equilibrium (SPE) are relaxed by equipping players with imperfect abilities of game theoretical reasoning. Delays and any possible divisions of a cake, not limited to SPE, are therefore possible. Having used a co-evolution adaptive system in which bargaining players learn "how to bargain from experiences" while competing against each others, we study experimentally how the discount factors affect the bargaining outcomes and compare those with the theoretical SPE. The findings reveal that the evolutionary computation approach is a convincingly complementary and approximating tool potentially able to tackle bargaining problems which would require excessive efforts if approached by traditional game-theoretic methods.

We start with brief descriptions to these methods and then focus on applying an evolutionary algorithm to an infinite extensive-form game, Alternating-offers bargaining problems.

1.1 Game theory

Game theory is a branch of mathematics that uses models to study interactions with formalized incentive structures ("games"). Game theory is important to many fields, including economics, biology, politics and computer science.

1.2 Analytical method

Von Neumann and Morgenstern first formalizes two-person zero-sum games and presents their theoretical optimal solutions for ideally rational players by equilibrating through mathematical reasoning ([Von Neumann & Morgenstern 1944]).

Game theorists solve games, assuming that every involver has "Perfect" rationality as an 'economic man' who typically has complete information relevant to problems, full computing capacity and well-defined and stable system of preferences. Rational players know all involvers are rational and know the rules of the game ([Simon 1955]).

[Nash 1950] formulates Nash Equilibrium for multiplayer games. Complex problems probably have multiple Nash Equilibriums. How to select equilibrium becomes a problem. Theorists have proposed different definitions of rationality to eliminate some equilibrium in order to refine the Nash equilibrium.

Game-theoretic analysis normally requires substantial time and costs. Theorists spend years to present a particular equilibrium for a particular situation. Substantial efforts may be required to find equilibriums for slightly modified situations.

1.3 Behavioral method

Some psychologists, along with social scientists and experimental economists collect data from human answers to questionnaires and competitions ([Simon 1982], [Barkow et al 1992], [Kagel & Roth 1995]). They observe and analyze *actual* human behaviors and try to explain why in some (simple) cases, people learn to perform better, *as if* they know theoretical equilibriums. In some other (often complicated) situations, people give intuitively reasonable responses, not using rational choices.

1.4 Evolutionary game theory

Maynard Smith and Price (1973) initiated *Evolutionarily Stable Strategy* (ESS) which is the most influential work since the Nash Equilibrium in Game Theory. Unfortunately, like traditional game-analytical methods, ESS does not explain *how* a population adapts to such a stable strategy [Weibull 1995]. Using ESS theory, one can check whether a strategy is robust to continually evolutionary pressures.

In the real world however, an ESS may not dominate in a population during a certain period of time, due to strong stochastic components emerging in evolutionary process. [Fogel et al. 1997] and [Fogel et al. 1998] show that "even in simple games, ESSs may not be stable under conditions that are pertinent in the real-world, such as finite population size and culling selection. Under proportional selection, large finite populations may tend to vary around an ESS, but large can be on the order of 5000 or more individuals in a population."

1.5 Evolutionary Algorithm Simulations

Evolutionary Algorithms refer to a class of algorithms which are inspired by natural evolution. Related methods to this work are Genetic Programming (GP) [Koza 1992] [Langdon & Poli 2001] and Genetic Algorithms (GA) [Holland 1975]. Evolution Algorithms are different from ESS, although both are rooted in evolutionary biology. To use an EA approach to solve games can be regarded as a way of simulation, which do not necessarily converge to game-theoretic equilibriums or ESS.

[Axelrod 1987] studies normal-form repeated games. His GA experimental results and [Miller 1996]'s results coincide with some reciprocity phenomena shown in human entries of Iterated Prisoners Dilemma (IPD) competitions. [Koza 1992] investigates a finite, extensive-form and no-repeated games with complete and perfect information, for which he finds Subgame Perfect Equilibrium using Genetic Programming (GP).

In this work, we attempt to employ Evolutionary Algorithm to study an infinite extensive-form two-person game with complete and perfect information: *Basic Alternating-Offer Bargaining Problem* (BAOBP), or Rubinstein Bargaining Problem [Rubinstein 1982] whose *Subgame Perfect Equilibrium* (SPE) is known. We start by introducing the BAOBP and its SPE. EA framework and a co-evolving system for BAOBP are developed, after which experimental outcomes are analyzed. Conclusions and future work will be given in the end.

2 Bargaining Problems

Bargaining problems study a class of situations where participants have common interests but conflict over how to divide the interest among them. Participants try to achieve agreements through negotiation. [Nash 1950] formulates the Nash Bargaining Problem and [Rubinstein 1982] models and solves the Basic Alternating-Offer Bargaining Problem. Based on these, other researchers study more complex bargaining situations.

2.1 Alternating-Offer Bargaining Problem

BAOBP describes a bargaining scenario in which the participant A makes an offer or count-offers to the player B on dividing a cake $\pi = 1$ at time of 0, 2, 4, 6, ... B makes a counter offer at time 1, 3, 5, 7, ... The bargaining process ends once an offer or a counter-offer is immediately accepted by the other player. A proposal on division by the player i is x_i for himself and $x_j = 1 - x_i$ for the other j. Player i's discount factor δ_i is his bargaining cost per time interval, $\delta_i \equiv e^{-r_i}$ where r_i is the player i's discount rate. The payoff gained by player i who has a share of x_i from the agreement, reached at time t is determined by the payoff function: $x_i \delta_i^t$.

2.2 Assumptions and Subgame Perfect Equilibrium

[Muthoo 1999] characterizes solutions to BAOBP problem by satisfying two properties: "no delay" and "stationarity". No delay means that "whenever a player has to make an offer, her equilibrium offer is accepted by the other player". Stationarity requires "in equilibrium, a player makes the same offer whenever she has to make an offer". Theorists mathematically analyze Subgame Perfect Equilibrium under such strong assumptions, in which players should offer nothing other than the perfect equilibrium partition and for sure will be accepted at time 0. Partitions are guaranteed before a bargain even starts, given the discounts factors.

The unique equilibrium taken as the Game-theoretic formula solution of this game is a Subgame Perfect Equilibrium in which the first player A obtains:

$$x_A^* = \frac{1 - \delta_B}{1 - \delta_A \delta_B}$$

and the second player B gets:

$$x_B^* = 1 - x_A^*$$

Technical treatments and proofs are available in [Rubinstein 1982], [Muthoo 1999], and [Bierman 1998].

3 Evolutionary Algorithms

Evolutionary algorithms are a population-based improvement mechanism. Individuals are selected based on their performance (fitness). Better individuals have higher probability to be selected as "raw material" to breed new offspring for the forthcoming generation. The offspring are created by the genetic operators (crossover and mutation) on the "raw" genetic material. Evolution pushes individuals (more specifically, the genetic materials) to continue improving their adaptation to the environments or objectives in order to survive. The improvement of individuals illustrates the process of acquiring behavior patterns adaptive to the environments.

In many applications, Evolutionary Algorithms are used as stochastic search methods, which are proposed to produce near-optimal solutions to a given problem. Given the size of the search space (depending on how strategies are represented), exhaustive search is normally impractical. An efficient approach able to search acceptable strategies within a reasonable time is therefore needed. EA are chosen not only because they have succeeded in many other applications, but also because they are expected that the same mechanism is applicable to slightly modified scenarios.

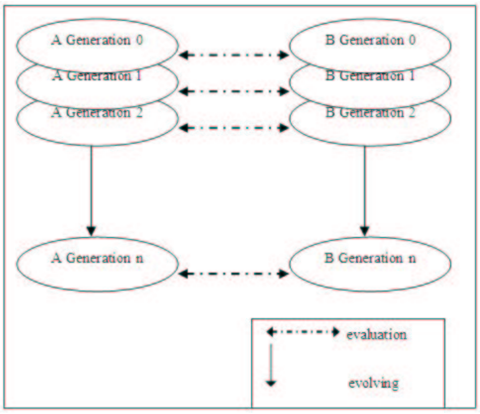

Figure 1: The diagram of Co-evolution

3.1 Co-evolution

Natural co-evolution is the mutual evolutionary influences between two species dependent on each another. The survival skills from co-evolving species in nature inspire scientists to borrow co-evolution principles to solve problems in which two elements are greatly interacting with each other. Figure 1 shows in an idealized two species situation, the species A and B are coexisting. One species's fitness is its current adaptation to the other species that is evolving simultaneously. Computer scientists have modeled [Schmitt 2004] and provided experimental outcomes to substantiate that co-evolution is practically good in some games. [Tsang & Li 2002] successfully developed a co-evolutionary system EDDIE/FGP which aids investors to seek dealing rules in financial markets.

4 Experiments

In this classical model of alternating-offers bargaining proposed by Rubinstein, the discount factors, specifying the respective costs subject to bargaining time for the two players, are the only elements that determine their bargaining powers. In this model, the theoretical SPE is unique and can be expressed analytically. However the assumptions of this model are too idealized hence other equilibriums and possibilities, which would arise in more realistic assumptions, are ruled out.

4.1 Assumptions to Players

Subsection 2.2 has provided the game-theoretic solution. The unique point prediction P.E.P analyzed by game theoretical method is a compelling one because it is difficult to see why perfectly rational players knowing all involvers are rational and knowing the rules of the game will do anything else. However, it is the assumptions that make applicators of the theory unlikely to believe provided that realistic players often lack the perfect rationality assumed. So, more realistic settings of assumptions have to be made in order to see whether strategies of players with Bounded Rationality converge to SPE. We hypothesize there are reciprocal interactions between players' behaviors, and players learn through trial-and-error experiences.

In our experiments, strategies have dynamic behaviors and can propose any division within the size of a cake rather than "stationarity"; bargaining procedure can last at most 10 time intervals (due to computational resources) instead of no delay; and players are allowed to use various strategies. Instead of assuming certain rationality of players, we hypothesize that

- Players in bargaining problems have a very low level of intelligence and are incapable of game-theoretic reasoning;
- The only goal of a player is to maximize his overall payoffs;
- Players learn through trial-and-error experiences over generations;
- There are reciprocal interactions between bargaining players' behaviors, like co-adapting organisms;
- A strategy has no ability of identifying opponents' strategies;
- A strategy has no memory of historical behaviors of the opponent's strategy in the undergoing bargaining;
- A strategy is unable to adjust its behaviors during a bargaining procedure. In other words, a strategy is a function without any parameters responding to its opponent's actions.

Using this set of assumptions avoids the difficulties of defining rationality.

4.2 Experimental Set-up

We build a two-population co-evolutionary system implemented by Genetic Programming (GP). Each player has his own population: a strategy pool consisting of candidate solutions. The strategies of the two players evolve at the same time. In this system, the objective of the strategies is to maximize the payoff from bargaining agreements. Well-performed strategies are chosen under the guideline of selection and undergo progressive modifications to be more competitive in forthcoming bargaining.

Genetic parameters are stated in Table 1.

Game parameters: 10 pairs of discount factors are chosen, which are shown in the Table 2. In theory, $0 < \delta_i < 1$. $\delta_i = 0.1$ means that a cake with a size 1 will shrink to be 0.1 after one time interval for the player i and $\delta_i = 0.9$ means that the cake will be 0.9 after one time interval. There are two examples of a low and a high discount factors.

Representation: An individual $g_i \in I$ is a genetic program in player i's population I. Its corresponding strategy

Table 1: GP parameters

Parameter	Value
Population Size	100
Number of Generations	300
Function Set	{ +, -, $\times$, $\div$ (Protected) }
Terminal Set	{1, -1, δ_A, δ_B }
Initial Max Depth	5
Initialization Method	Grow
Selection Method	Tournament
Crossover rate	0 to 0.1
Mutation rate	0.01 to 0.5
Maximum nodes of a GP program	50

is $s(g_i)$. In order to make the search space smaller, currently we evolve only g_i. A time-dependent strategy of player i is $s(g_i) = g_i \times (1 - r_i)^t$, where t is time, an non-negative integer.

Fitness Function: A strategy's performance highly depends on other strategies whom it meets. The design of using a group of fixed representative strategies as the fitness assessment has a risk that evolution may exploit the weaknesses of the pre-defined representatives, but perform poorly against others. So the fitness of a strategy should be based on its performance against the opponent's co-evolving strategies at the same evolutionary time. In other words, for this bargaining problem, the relative fitness [Koza 1992] assessment is a fair choice. Game Fitness of a strategy $s(g_i)$, denoted by $GF(s(g_i))$ is defined as the average payoff of $s(g_i)$ gained from agreements with strategies in the opponent's population J which has a set of n number of bargaining strategies, $j \in J$:

$$GF(s(g_i)) = \frac{\sum_{j \in J} p_{s(g_i) \to s(g_j)}}{n}$$

where $p_{s(g_i) \to s(g_j)}$ is the payoff gained by $s(g_i)$ from an agreement with $s(g_j)$ which receives $p_{s(g_j) \to s(g_i)}$. An Incentive Method to handle constraints is used in defining the fitness function for all the individuals in the populations. Detailed designs of the fitness function are described in [Tsang & Jin 2004].

4.3 Observations

We have executed 100 runs with different random seeds for each game's setting chosen. For each game's setting, the average of shares x_A from final agreements made by the best-of-generation individuals from both populations at the 300^{th} generation is shown in the table 2. We observe that after 300 generations, the 100 x_As cluster around SPE, having minority of exceptions found. To our hypothesis that SPE is the same as the mean of our experimental shares, a t-test shows the t Critical value two-tail is 2.2621, larger than the t statistics value 1.3011. So our hypothesis is accepted at the 95% confidence level. Many experiments,

Table 2: The means of the shares x_As obtained by the best-of-generation individual in the population A at the 300^{th} generation

Discount Factors	SPE x_A^*	Experimental Average x_A	x_As' Deviation
(0.1, 0.4)	0.6250	0.9101	0.0117
(0.4, 0.1)	0.9375	0.9991	0.0054
(0.4, 0.4)	0.7143	0.8973	0.0247
(0.4, 0.6)	0.5263	0.5090	0.0096
(0.4, 0.9)	0.1563	0.1469	0.1467
(0.5, 0.5)	0.6667	0.6745	0.0271
(0.9, 0.4)	0.9375	0.9107	0.0106
(0.9, 0.6)	0.8696	0.8000	0.1419
(0.9, 0.9)	0.5263	0.5065	0.1097
(0.9, 0.99)	0.0917	0.1474	0.1023

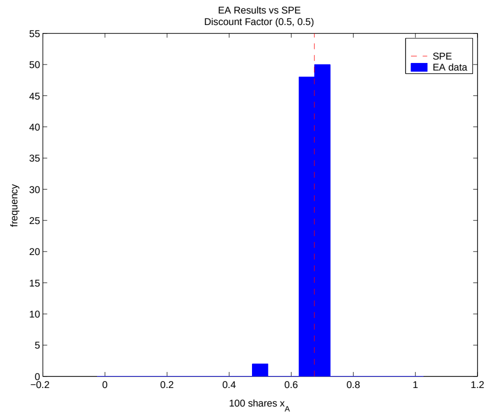

Figure 2: The distribution of 100 runs x_A at 300^{th} generations: δ_A = 0.5 and δ_B = 0.5. The vertical line x = 0.6667. is the SPE x_A^*.

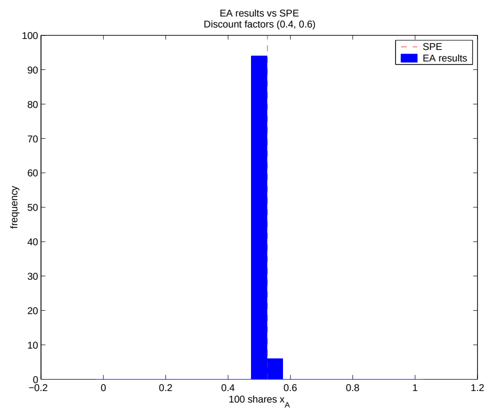

Figure 3: The distribution of 100 runs x_A at 300^{th} generations: δ_A = 0.4 and δ_B = 0.6. The vertical line x = 0.5263. is the SPE x_A^*.

Table 3: Bargaining Time for agreements made by the best-of-generation individuals from the two populations at the 300^{th} generation

Discount Factors	Average Time
(0.1, 0.4)	0.0000
(0.4, 0.1)	0.0000
(0.4, 0.4)	0.0000
(0.4, 0.6)	0.0000
(0.4, 0.9)	0.2121
(0.5, 0.5)	0.0000
(0.9, 0.4)	0.0100
(0.9, 0.6)	0.4700
(0.9, 0.9)	3.8500
(0.9, 0.99)	5.6100

the SPEs are within the distributions of experimental outcomes for example in Figure 2 and Figure 3. Therefore it is very likely that co-evolutionary might generate exact solutions as the theoretical ones, at a certain degree of precision. For game settings with extreme low or high discount factors, experimental results are far from the SPE predictions. In our experiments, extreme bargaining parameters refer to the sets of discount factors: (0.1, 0.4), (0.4, 0.4) and (0.9, 0.99). [Bragt et al. 2002] simulates the bargaining by a multi-agent evolving system that is implemented by real number-coded Genetic Algorithms. They have found similar results although their experiments only test the situations when the $\delta_A = 0.6$ or $\delta_A = 0.3$.

All experimental results clearly demonstrate the influence of discount factors upon bargaining powers: the player with higher discount factor, comparative to his opponent, attains a larger portion of cake. If both players have the same discount factors, the first player receives a larger part of division. This discovering is consistent with the analysis by bargaining theory [Muthoo 1999].

The discount factors also determine the negotiation time required for settlements (Table 3). Not all bargains reach an agreement at the time $t = 0$. Delays ($t > 0$) emerge as a consequence of players' preferences to higher payoff and expectations that higher payoff will obtain in future, by (mainly) those players who have high discount factor; i.e. relatively patient players. Impatient players, on the other hand, are eager to agree as soon as possible to avoid such delays due to relatively higher costs per time interval they should pay than that of patient players should. Any delay ($t > 0$) is costly to both of the two players. Thus the total sum of payoff they get is less than the size of cake, which means the cake is not divided efficiently.

Moreover, the EA approach opens a window to show the process of artificial evolving over time, which neither game-theoretic solution nor ESS can provide. In Figure 4 and 5, two players' discount factors are (δ_A, δ_B) = (0.9, 0.4). The horizontal line of 0.9375 and the horizontal line of 0.0625 are SPE for player A and B respectively. The values of g_i (labeled as GPT), the shares from agreements

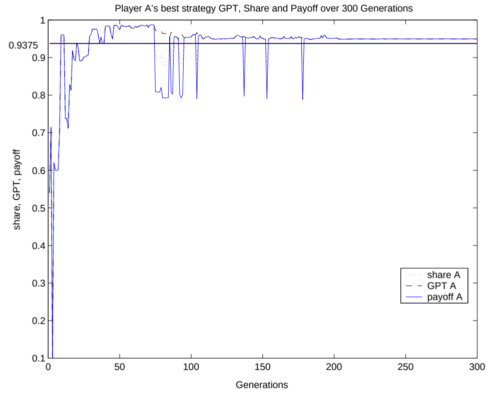

Figure 4: The best-of-generation individual of Population A over 300 generations. (δ_A =0.9 and δ_B = 0.4)

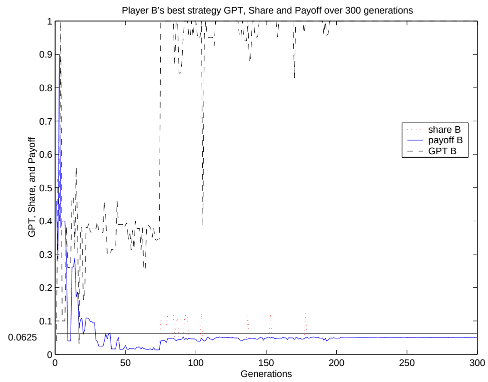

Figure 5: The best-of-generation individual of Population B over 300 generations. (δ_A =0.9 and δ_B = 0.4)

and the payoffs of the corresponding shares of the best-of-generation strategies of every generation are displayed. Strategies continually update themselves over time to co-adapt each other and come up with using relatively stable strategies, measured by fitness. We comment now on modifications of strategies' behaviors in this typical run. In the initial population strategies are generated randomly and for both players this means offering a deal of roughly 50% of the cake in average. Soon after, the player A learns that he can obtain more, because for him, delay is less costly as he has a higher discount factor than his opponent. He changes his first offers as much as he can in order to maximize his payoff. Finally, he approaches a value relatively close to the theoretical perfect equilibrium. Player B learns that she has to secure an agreement as soon as possible because it is not worthy for her to wait. Thus players finally reach at the point where both of them are willing to agree at time 0 when no bargaining costs occur. This fits nicely with the theoretical explanation.

5 Conclusions

Our objective was to compare the EA results with the Subgame Perfect Equilibrium solutions. The experimental observations show that co-evolutionary algorithms ideally provides approximate SPE solutions, excluding the game's settings having very high and/or very low discount factors. Although co-evolution has not always evolved identical solutions to SPE in our experiments, it has given approximate alternatives. Moreover, this approach has provided interesting information concerning the influences of discount factors on bargaining time, the divisions of the cake and the evolving process. Compared with other research methods mentioned in the section 1, an EA approach has particular advantages: on much lower costs over human subjects' experiments; on less human intelligence over economics theorists and on its reusability and modifiability for complicated bargaining situations without knowledge of game-theoretic solutions such as variants of basic alternating-offer bargaining problems and Incomplete information bargaining problems.

In this work, we have only studied one bargaining problem, and emphasize that considerably more work will be required to determine the general utility of EAs in the problem domain. In future, we also plan to compare the EA results with actual behaviors by human-subject experiments.

Acknowledgement

This research is partly sponsored by University of Essex Research Fund. We benefited greatly from conversations with Abhinay Muthoo and Riccardo Poli. Alberto Moraglio, Yossi Borenstein and Vikentia Provizionatou have given important feedbacks on the draft of this paper. Mathias Kern helped with the use of LaTeX. Anonymous reviewer's comments are very helpful for the revision. Tim Gosling helped improve the quality of the final submission.

Bibliography

[Axelrod 1987] Axelrod, R. (1987) The evolution of strategies in the iterated prisoner's dilemma. In Lawrence Davis (Ed.), Genetic Algorithms and Simulated Annealing. Morgan Kaufmann.

[Barkow et al 1992] Barkow, J., Cosmides,L., Tooby, J. edited (1992) The Adapted mind : evolutionary psychology and the generation of culture, Oxford University Press.

[Bierman 1998] Bierman, H. and Fernandes, L. (1998) Game Theory with Economic Applications, Addison-Wesley, New York.

[Bragt et al. 2002] van Bragt, D., Gerding,E., and La Poutré, J. (2002) Equilibrium Selection in Alternating-Offers Bargaining Models: The Evolutionary Computing Approach, Electronic Journal of Evolutionary Modeling and Economic Dynamics (e-JEMED) 2.

[Fogel et al. 1997] Fogel, D., Fogel, G., Andrews, P., (1997) On the instability of evolutionary stable strategies, BioSystems 44, 135-152.

[Fogel et al. 1998] Fogel, G., Andrews, P., Fogel, D. (1998) On the instability of evolutionary stable strategies in small populations, Ecological Modelling 109, 283-294.

[Gibbons 1992] Gibbons, R. (1992) A Primer in Game Theory, Harvester Wheatsheaf.

[Holland 1975] Holland, J. (1992) Adaptation in Natural and Artificial Systems, the second edition, MIT Press.

[Koza 1992] Koza, J. (1992) Genetic Programming: On the Programming of Computers by Means of Natural Selection. The MIT Press.

[Langdon & Poli 2001] Langdon, W.B. and Poli, R. (2001) Foundations of genetic programming, Springer.

[Miller 1996] Miller, J. (1996) The Co-evolution of Automata in the Repeated Prisoner's Dilemma. Journal of Economic Behavior and Organization, 29(1), 87-112.

[Muthoo 1999] Muthoo, A. (1999) Bargaining Theory with Applications, Cambridge University Press.

[Nash 1950] Nash, J. (1950) The Bargaining Problem, Econometrica, 18: 155-162

[Rubinstein 1982] Rubinstein, A. (1982) Perfect Equilibrium in a Bargaining Model, Econometrica, 50: 97-110.

[Schmitt 2004] Schmitt, L. (2004) Classification with Scaled Genetic Algorithms in a Coevolutionary Setting. GECCO (2): 138-149

[Simon 1955] Simon, H. (1955) A Behavioral Model of Rational Choice, The Quarterly Journal of Economics, Vol. 69, NO.1, 99-118

[Kagel & Roth 1995] Kagel, J. and Roth, A editors (1995) The Handbook of Experimental Economics, Princeton University Press.

[Maynard Smith 1982] Maynard Smith, J. (1982) Evolution and the theory of games, Cambridge University Press.

[Nawa 2003] Nawa,N., Shimohara,K., and Katai, O. (2003) Diversity in Time Preferences and Fairness in Bargaining Games Played by Evolutionary Agents, The Second International Conference on Computational Intelligence, Robotics and Autonomous Systems.

[Simon 1982] Simon, H. (1982): Models of bounded rationality, Cambridge, Mass, MIT Press.

[Tsang & Jin 2004] Tsang, E., Jin, N. (2004) Incentive Method for Constraint Handling used with Evolutionary Algorithms, Technical report CSM-417, Department of Computer Science, University of Essex.

[Tsang & Li 2002] Tsang, E., Li, J. (2002) EDDIE for financial forecasting, in S-H. Chen (ed.), Genetic Algorithms and Programming in Computational Finance, Kluwer Series in Computational Finance.

[Von Neumann & Morgenstern 1944] Neumann, J. and Morgenstern, O. (1944) Theory of games and economic behavior, Princeton University Press.

[Weibull 1995] Weibull, J. (1995) Evolutionary game theory, Cambridge, Mass. : MIT Press.

[Wikipedia] Wikipedia, the free encyclopedia on line: http://en.wikipedia.org/