On using ZCS in a Simulated Continuous Double-Auction Market

Larry Bull
Intelligent Computer Systems Centre
Faculty of Computer Studies and Mathematics
University of the West of England
Bristol BS16 1QY, U.K.
larry@ics.uwe.ac.uk

Abstract

This paper presents results from on-going investigations into the performance of the Michigan-style classifier system in a complex multi-agent environment. Using a simplified model of a continuous double-auction market place the use of ZCS as an adaptive economic trading agent is examined. It is shown that a number of small changes to the basic system greatly improves its performance, resulting in improvements in the overall efficiency of the market. It is also shown that the role of the rule-discovery component of the classifier system is particularly critical in such a closely-coupled multi-agent environment.

1 INTRODUCTION

As evolutionary computing techniques are applied to multi-agent environments, new issues arise along with new view points on traditional areas of the use of such approaches. In this paper the use of a Michigan-style classifier system [Holland & Reitman 1978], Wilson's Zeroth-level classifier system (ZCS)[Wilson 1994], in a complex multi-agent environment, that of a simulated continuous double-auction market, is investigated.

Continuous double-auction markets are a type of economic trading forum in which traders make bids and offers for goods, based on other traders' bids and offers. In this paper a simulated continuous double-auction is presented in which traders are each represented by a ZCS and hence the classifier systems must learn suitable strategies to buy and sell goods effectively. In this paper it is shown that ZCS agents are able to learn effective trading strategies over time and that the role of the genetic algorithm (GA) [Holland 1975] can be particularly significant in such systems. This work has been inspired, in part, by the growing interest in the use of evolutionary computing techniques in general to study economics using computer simulations [e.g. Chattoe 1994], for which classifier systems appear particularly well suited [Holland & Miller 1991].

The paper is arranged as follows: the next section briefly introduces the ZCS architecture used here. Section 3 reviews previous work using classifier systems in multi-agent environments in general and section 4 reviews previous work on their use within computational economic modeling. Section 5 describes the simulated continuous double-auction market used for this work and section 6 presents results from its use with multiple ZCS traders.

2 ZCS

ZCS is a Michigan-style Classifier System without internal memory, where the rule-base consists of a number (N) of condition/action rules in which the condition is a string of characters from the usual ternary alphabet {0,1,#} and the action is represented by a binary string. Associated with each rule is a strength scalar which acts as an indication of the perceived utility of that rule within the system. This strength of each rule is initialized to a predetermined value termed S_0.

Reinforcement in ZCS consists of redistributing strength between subsequent "action sets", or the matched rules from the previous time step which asserted the chosen output or "action." A fixed fraction (β) of the strength of each member of the action set ([A]) at each time-step is placed in a "common bucket." A record is kept of the previous action set $[A]_{-1}$ and if this is not empty then the members of this action set each receive an equal share of the contents of the current bucket, once this has been reduced by a predetermined discount factor (γ). If a reward is received from the environment then a fixed fraction (β) of this value is distributed evenly amongst the members of [A]. Finally, a tax (τ) is imposed on all matched rules that do not belong to [A] on each time-step in order to encourage exploitation of the stronger classifiers. Hence this is different from the traditional "Bucket-brigade" algorithm [Holland et al. 1986] and is known [Wilson 1994] to be very similar to Watkin's Q-learning [1989] reinforcement algorithm.

ZCS employs two discovery mechanisms, a panmictic GA (altered later here) and a covering operator. On each time-step there is a probability *p* of GA invocation. When called, the GA uses roulette wheel selection to determine two parent rules based on strength. In this paper, as in [Cliff & Ross 1994], one offspring is produced via mutation (probability ν) and crossover (single point with probability χ). The parents then donate a third of their strengths to their offspring who replaces an existing member of the rule-base. The deleted rule is chosen using roulette wheel selection based on the reciprocal of rule strength. If on some time-step, no rules match or all matched rules have a combined strength of less than ϕ times the rule-base average, then a covering operator is invoked exactly as in [Wilson 1994].

The default parameters presented for ZCS, and unless otherwise stated for this paper, are: N = 400, S_0=20, β = 0.2, γ = 0.71, τ = 0.1, χ = 0.5, ν = 0.001, p = 0.25, ϕ = 0.5

Thus ZCS represents a "basic classifier system for reinforcement learning that retains much of Holland's original framework while simplifying it so as to increase understandability and performance" [Wilson 1994]. For this reason the ZCS architecture has been chosen to examine the basic behaviour of classifier systems in a complex multi-agent environment, as opposed to the more sophisticated systems of Holland et al. [1986] or Wilson's XCS [Wilson 1995] for example.

The reader is referred to [Wilson 1994] for full details of ZCS.

3 CLASSIFIER SYSTEMS IN MULTI-AGENT ENVIRONMENTS

A small number of investigators have examined the use of classifier systems in multi-agent environments. Bull et al. [e.g. 1993, 1995] describe the use of Pittsburgh-style [Smith 1980] classifier systems for the control of simulated robots, where each wheel/leg is represented by a separate system. Carse et al. [e.g. 1995] have used fuzzy Pittsburgh-style classifier systems for routing at each node of a telecommunications network. Pittsburgh-style systems which also use reinforcement learning have been coevolved by Potter et al. [1995], where an agent is represented by a number of classifier systems and a speciation-like process is included to improve performance. Multiple Michigan-style classifier systems have been used by Dorigo and Schnepf [e.g. 1992] to control an autonomous robot and by Seredynski et al. [1995] to examine the use of local reward sharing in a simple iterated game.

4 CLASSIFIER SYSTEMS AS ADAPTIVE ECONOMIC AGENTS

Apart from the work described in the previous section, the only other known body of work examining the use of Michigan-style classifier systems in multi-agent environments exists in the field of computational economics. After [Arthur 1990] and [Holland & Miller 1991] a number of researchers have used classifier systems to represent traders in artificial markets. Marimon et al. [1990] use classifier agents exchanging and consuming goods to examine the emergence of equilibria in a well-known triangular market. Palmer et al. [1994] describe the use of classifier systems to simulate agents creating portfolios, by predicting the value of a stock. Dwormann [e.g. 1994] has investigated coalition formation in a three-player game/market and Morengo and Tordjman [1996] used a classifier-based system to model belief formation in a market place. In this paper the use of ZCS in an artificial continuous double-auction (CDA) market is presented.

5 A SIMULATED CONTINUOUS DOUBLE-AUCTION MARKET

A continuous double-auction market consists of a number of buyers and sellers with reservation prices (known only to the agents themselves). A reservation price is the price below/above which a trader will not sell/buy. At any time during a trading session (day) buyers can make bids and sellers can make offers, with trades occurring when an agent's bid/offer ("shout") is accepted by another agent. Note that an agent's profit margin is defined by the difference between its reservation price and its current shout price. The London and New York stock exchanges use CDAs, and hence there is much interest in their behaviour [e.g. Friedman & Rust 1992]. The efficiency of such markets can be determined using Smith's [Smith 1992] convergence metric, α, defined as $\alpha = 100\sigma/P_e$, where P_e is the price at which the quantity demanded is equal to the quantity supplied - the equilibrium price - and σ is the standard deviation of trade prices around P_e. Hence the lower α, the more effective the traders and the more efficient the market place.

In the simulations used here, trades occur automatically when two current shout prices "overlap", e.g. when a given seller's shout is less than or equal to a given buyer's shout. On each trading day each agent has one unspecified commodity to buy (buyer) or sell (seller), with the same starting shout price/profit margin as at the end of the previous day. Agents are assigned their reservation prices by dividing the price range of the market equally between the number of agents. For example, in the experiments presented here there are three sellers and three buyers, all with prices in the range 75-325 units, giving an equilibrium price of 200 units

(Figure 1). A day's trading finishes when the remaining agents' reservation prices are known to not overlap (since no further trades are possible).

On any discrete time step, or round, within a trading day agents are told whether the last shout price was higher or lower than their current shout price, whether it was a bid or an offer, and whether it was accepted and a trade occurred.

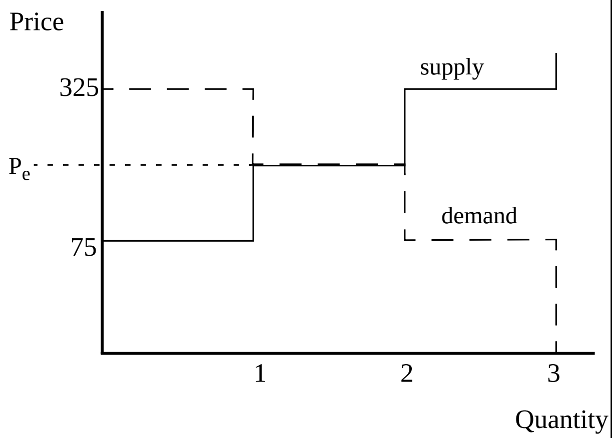

Figure 1: Showing how the reservation price of each agent is assigned from the possible price range of the CDA market, i.e. the supply (sellers) and demand (buyers) curves. There are three sellers, each with one unit to sell, and three buyers, each wanting to buy one unit, per day. The equilibrium price of the market is indicated, that is, the most efficient market trading price P_e is shown.

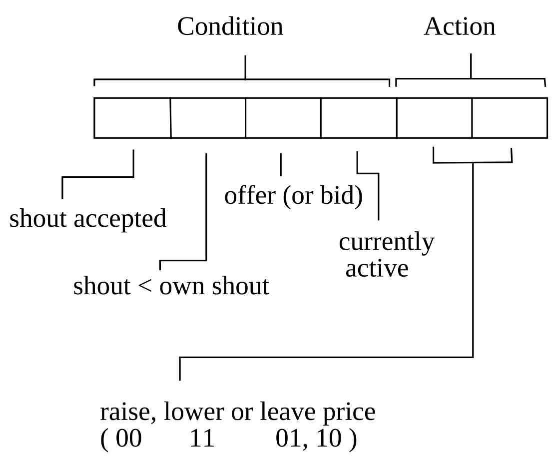

Figure 2: The structure of the classifier system rules.

The agents also know whether they have already traded on that day; inactive agents are able to "watch" the market. This information is generated by a central "pit" in the market, which is also responsible for spotting and handling trades between agents; a central process exists within the simulated market with which all agents interact (agents do not interact directly).

In these experiments a varying number of agents are represented by ZCS, with the aim of learning trading rules within the CDA. On each round, an agent receives the inputs described above, coded as a binary string and returns an action to raise, lower or maintain its current shout price (Figure 2). Profit margins (μ) on reservation prices (λ) are initially assigned at random, though shout prices (ρ) remain within market limits, and thereafter are adjusted in steps defined by: $\mu(t+1) = (\rho(t)+\Gamma(t))/\ \lambda - 1$, where Γ is a damping factor (Widrow-Hoff delta rule with the last shout price as the target - see [Cliff & Bruten 1998] for full details). Agents are rewarded only when they trade (1000μ).

Hence the number of ZCS outputs (shouts) needed before reward is received varies and each agent is highly dependent upon the others in the market; the CDA represents a fairly complex multi-agent environment.

6 RESULTS

The performance in the simulated CDA of three ZCS seller agents trading with three buyer agents which use the "ZIP" strategy (Figure 3) has been investigated. Agents using the ZIP trading strategy have been shown to quickly create an optimally efficient CDA market [Cliff & Bruten 1998].

if (the last shout was accepted at price *q*)
then
  1. if any buyer for which $\rho > q$ should raise its profit margin
  2. if (the last shout was an offer)
     then
       1. any active buyer for which $\rho < q$ should lower its margin
else
  1. if (the last shout was a bid)
     then
       1. any active buyer for which $\rho < q$ should lower its margin

Figure 3: The rules for a "Zero Intelligence Plus", ZIP, (buyer) trader.

Figure 4(a) shows results from a typical run (from 50), over 5000 days in which the ZCS had the same parameters as those used in [Wilson 1994] (see section 2). Figure 4(b) shows the results from a typical run in which a number of slight changes were made to the basic ZCS system. Previously, it has been shown [Bull 1998] that reducing the learning rate (β=0.1, τ = 0.05), reducing the rule-discovery rate (p=0.1), using a simple niche GA (after [Booker 1985]) and increasing the action-selection pressure (by squaring rule bids) improves ZCS's performance in multi-agent environments. Here the market is found to be more efficient, i.e. α is lower and for longer periods, with sellers using the slightly altered version of ZCS; only 11 of the 50 runs with the parameters in [Wilson 1994] showed any kind of stasis, whereas 34 of the 50 runs with the modified parameters gave significant periods of efficient trading.

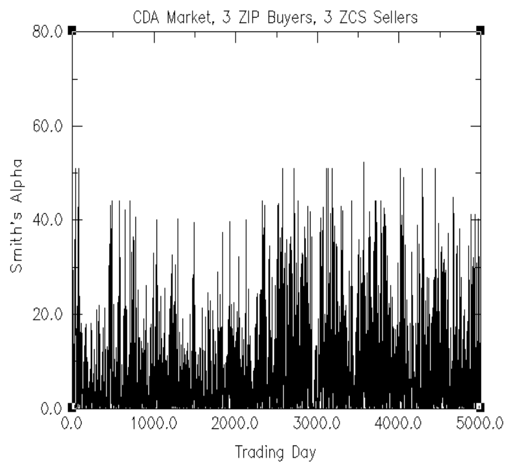

Figure 4(a): Typical performance of basic ZCS in the artificial CDA described in the text.

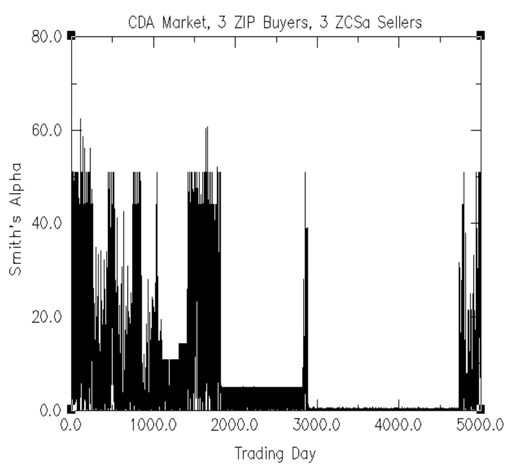

Figure 4(b): Typical performance of the altered ZCS in the artficial CDA market.

This has been explained as causing a reduction in the amount of exploration by each agent, creating a more stable environment in which the agents adapt [Bull 1998] (based on the seminal work of Kauffman and Johnsen [1991]).

It can also be seen in Figure 4(b) that, even with less exploration by the ZCS agents, the typical progression over time of the market with the altered ZCS traders contains periods of stasis, followed by bursts of change; punctuated equilibria are seen [Eldridge & Gould 1972]. It is also noted that, even in these improved markets, near-optimal periods of trading can be lost, as seen around day 4800.

Figure 5 shows a shorter run in which the punctuated characteristic is very clear. After an initial period of price adjustment, the market settles to trading near the optimum

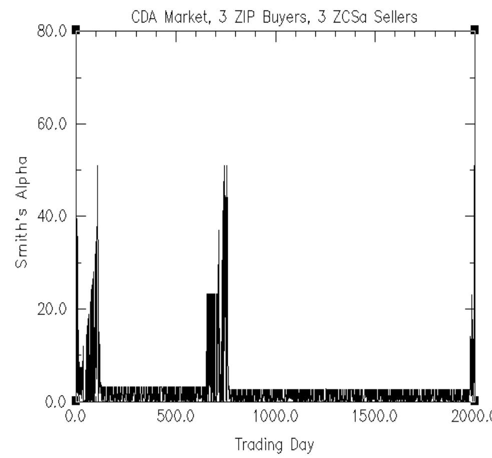

Figure 5: Progress of a run in which the cause of price fluctuations is investigated.

price, around day 200. However, around day 680 the market suffers from a period of inefficient trading, before settling down again to a low α around day 800. The rules of all the ZCS sellers were saved every 10 days during this run. Analysis shows that during the first few days sellers have rules which raise/lower their profit margins when the current shout price is higher/lower (akin to the ZIP strategy). By day 200 the sellers have learnt to leave their prices unaltered, under all conditions; the rule bases contain quite general rules with actions which do not adjust profit margins. Indeed, to some extent, the conditions of rules are free to "drift" in their space of possibilities so long as their actions do not alter the shout price. This stability changes around day 640: one seller has a new rule which means that it lowers its price when an offer lower than its own current shout price is made. Inspection shows that this rule has been created by the GA through recombination and the mutation of an action bit. Therefore, this rule causes the seller to go back to altering its shout price, which causes the other agents to alter their prices in response, resulting in a period of fluctuation.

Inspection of the same agent's rules at day 800 reveals a return (roughly) to the previous set of generalists which leave the shout price unaltered. That is, because of the use of strength/fitness inheritance, a different (mutated offspring) rule is chosen by the action selection mechanism. In this case the rule leads to less payoff than its parents and so its inherited strength is decreased by the reinforcement component. Eventually the mutant's strength drops sufficiently low in comparison to that of its parents that they again become selected as actions for the agent; the reinforcement learning process recovers from the detrimental exploration of the GA.

Another period of price fluctuation can be seen to be occurring at the end of the run, around day 2000. The periodic patterns seen in the market are due to different agents trading with different individuals on each day. A number of similar runs were examined with similar phenomena found.

Although this process causes the market to lose equilibria, from Figure 4(b) it can also be seen as beneficial to the market as a whole since the α of the new equilibrium is often lower than that previously obtained, e.g. compare days 1500, 2500 and 3500; GA generated rules can improve trader behaviour.

The same experiments have been run in which *all* traders are represented by ZCS agents. Here experiments were left to run for 30,000 trading days to allow for the extra learning involved. Figure 6 shows how the dynamics of these systems are different from those above. It has been found that, with the improved parameter settings, in a large number of runs (19/20) the market settles into an attractor in which the average α is high (around 50), with pronounced peaks and dips; in twenty runs two markets found low α equilibria for a significant period of time. The simulation in Figure 6 shows two "snap shots" from a run in which a low equilibrium was found. The market can be seen to enter the common, less efficient state around day 250, but around day 23,000 the market settles into a near optimal state for around 2000 days, before returning to the less efficient state.

Again, runs have been carried out in which agents' rules were periodically saved. It has been found that in these runs, as with the previous ones, agents eventually contain rules which do not alter their shout prices over all conditions. The difference is that here some agents seem to gain an initial advantage and effectively "dictate" to the rest of the (small) market. It appears that some agents, e.g. the sellers, quickly learn to alter their shout prices in order to be involved in a trade, since that is the only way credit can be received. However the other agents, e.g. the buyers, do not learn to adjust their price but end up in a trade anyway. Hence the sellers make very small profits and the buyers very large profits, resulting in the whole market being well away from an efficient equilibrium. This market configuration then gets reinforced because even if the sellers try to raise the price, the buyers have no incentive to lose reward; initial stupidity can pay! Again, the dips in α are due to different combinations of agents trading (price distributions are large here).

It is suggested that low equilibrium markets are hard to find here because both sellers and buyers must almost simulta-

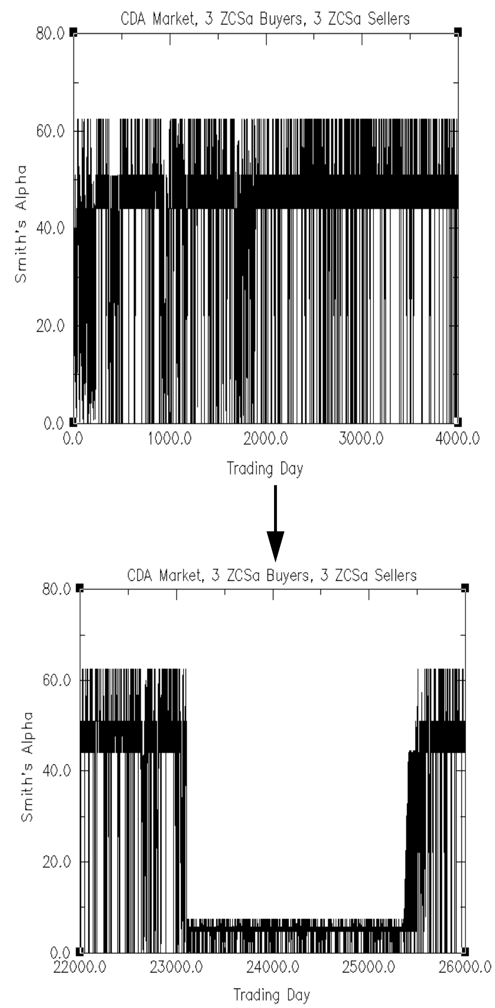

Figure 6: Performance of CDA market with all traders represented by ZCS agents. The first graph shows the common high α state and the second a temporary low state.

neously discover rules which will cause them to alter their profit margins.

Note that in the previous markets, a move to adjust the current trading price by a ZCS seller almost always causes the

ZIP buyers to alter their shout prices; the ZIP strategy does not consider agent profit.

7 CONCLUSIONS

This paper has presented results from on-going investigations into the performance of the Michigan-style classifier system in a complex multi-agent environment. Using a simplified model of a continuous double-auction market place the use of ZCS as an adaptive economic trading agent has been examined. It has been shown that a number of small changes which reduce the amount of exploration in ZCS greatly improves its performance in the system, resulting in improvements in the overall efficiency of the market.

It has also been shown that the role of the rule-discovery component of the classifier system is particularly critical in such a tightly coupled multi-agent environment. The use of fitness/strength inheritance allows less useful offspring rules to be selected as outputs, which has been shown to disrupt equilibria within the market. Equilibria are restored once the reinforcement learning component has altered such rules' strengths to that appropriate to their actions; over time, trading is punctuated with bursts of inefficient trading. It was also seen that offspring rules may improve trading strategies and so any new equilibrium may be more efficient than before, as expected. Note that this highlights a significant difference in the behaviour of a multi-agent system containing classifier systems to one containing many other forms of evolutionary computing and machine learning techniques. Typically, other techniques will eventually converge upon a single solution for each agent, in the context of the other agents, creating a system equilibrium from which it is unlikely agents will deviate; most other techniques do not include a source of agent action utility "noise" allowing them to move significantly from a given optimum (see also [Lindgren 1991] for an Iterated Prisoner's Dilemma [Axlerod 1987] model exhibiting similar behaviour due to the inclusion of particular search operators).

In the closely related extension of [Palmer et al. 1994], it has been shown that an artificial stock market can exhibit two types of useful behaviour, depending on the rate at which the GA is used [Arthur et al. 1997]: when the GA rate is low, the market behaves as predicted by theoretical models with stable, near optimal trading; and with a higher rate, the market behaves as real markets do, with periods of fluctuations and instabilities. Arthur et al. used a version of Holland's original classifier system, preprogrammed with useful rules (presumably because of the model's complexity), but using a deterministic "the best" action selection policy, i.e. a policy with low exploration (see [Bull 1998] for an examination of such a strategy's performance in classifier systems starting with random rules in a multi-agent system). That is, *rule discovery appears to be a very significant aspect of tightly coupled systems in general*.

Based on these findings the use of a triggered GA [Booker 1989] may provide significant benefits in such systems (see also [Wilson 1995]).

The results from this simulated economic model are also relevant to the field of market-based control [e.g. Miller & Drexler 1988], in which free-market mechanisms are used to control distributed computer systems. Such techniques have been used with designed agents to assign network bandwidth [e.g. Miller et al. 1996] and computer memory [e.g. Harty & Cheriton 1996], to control air-conditioning systems [Clearwater et al. 1996], etc. The results here suggest that it may also be possible to use adaptive agents in such systems; the above experiments containing much larger numbers of agents are now being carried out.

Further enhancements to Michigan-style classifier systems for use in this multi-agent environment are currently being investigated, for example it has been suggested that the use of memory in such non-Markov environments can be beneficial (mentioned for ZCS in [Wilson 1994] and implemented in [Cliff & Ross 1994], see also [Tomlinson & Bull 1998] for another mechanism). This is now being investigated, along with the use of other forms of classifier system such as XCS [Wilson 1995].

8 ACKNOWLEDGEMENTS

This work was supported by Hewlett Packard Laboratories, Bristol. Thanks to Janet Bruten and the Agent Technology Group for the use of their software.

9 REFERENCES

Arthur W B (1990), "A Learning Algorithm that Replicates Human Learning", Technical Report 90-026, Santa Fe Institute.

Arthur W B, Holland J H, Palmer R & Tayler P (1997), "Asset Pricing Under Endogenous Expectations in an Artificial Stock Market", in W B Arthur, S Durlauf & D Lane (eds.) *The Economy as an Evolving Complex System II*, Addison-Wesley.

Axlerod R (1987), "The Evolution of Strategies in the Iterated Prisoner's Dilemma," in L Davis (ed.) *Genetic Algorithms and Simulated Annealing*, Pittman, pp32-42.

Booker L (1985), "Improving the Performance of Genetic Algorithms in Classifier Systems", in J J Grefenstette (ed.) *Proceedings of the First International Conference on Genetic Algorithms and their Application*s, Lawrence Erlbaum, pp80-93.

Booker L (1989), "Triggered Rule Discovery in Classifier Systems", in J D Schaffer (ed.) *Proceedings of the Third International Conference on Genetic Algorithms*, Morgan Kaufmann, pp265-275.

Bull L (1998), "On ZCS in Multi-Agent Environments" in A E Eiben, T Baeck, M Schoenauer & H-P Schwefel (eds) *Parallel Problem Solving from Nature - PPSN V*, Springer Verlag, pp471-480.

Bull L & Fogarty T C (1993), ""Coevolving Communicating Classifier Systems for Tracking", in R F Albrecht, C R Reeves & N C Steele (eds.) *Artificial Neural Networks and Genetic Algorithms*, Springer Verlag, pp522-527.

Bull L, Fogarty T C & Snaith M (1995), "Evolution in Multi-Agent Systems: Evolving Communicating Classifier Systems for Gait in a Quadrupedal Robot," in L J Eshelman (ed.) *Proceedings of the Sixth International Conference on Genetic Algorithms*, Morgan Kaufmann, pp382-388.

Carse B, Fogarty T C & Munro A (1995), "Adaptive Distributed Routing using Evolutionary Fuzzy Control", in L J Eshelman (ed.) *Proceedings of the Sixth International Conference on Genetic Algorithms*, Morgan Kaufmann, pp389-397.

Chattoe E (1994), "The use of Evolutionary Algorithms in Economics: Metaphors or Models of Social Interaction ?", in E Hillebrand & J Stender (eds.) *Multi-Agent Simulation and Artificial Life*, IOS, pp48-83.

Clearwater S, Costanza R, Dixon M & Schroeder B (1996), "Saving Energy Using Market-Based Control", in S Clearwater (ed.) *Market-Based Control: A Paradigm for Distributed Resource Allocation*, World Scientific, pp253-273.

Cliff D & Ross S (1994), "Adding Temporary Memory to ZCS", *Adaptive Behaviour* 3:101-150.

Cliff D & Bruten J (1998), "Less than Human: Simple Adaptive Trading Agents for CDA Markets", in *Proceedings of the 1998 IFAC/SCE/IFIP/IFORS Society for Computational Economics Symposium on Computation in Economics, Finance and Engineering: Economic Systems*.

Dorigo M & Schnepf U (1992), "Genetics-based Machine Learning and Behaviour-based Robotics: A New Synthesis", *IEEE Trans. on Sys. Man and Cybernetic*s 22(6):141-154.

Dworman G (1994), "Games Computers Play: Simulating Characteristic Function Game Playing Agents with Classifier Systems", in *Proceedings of the 1994 IEEE Conference on Evolutionary Computation*, IEEE, pp696-701.

Eldridge N & Gould S J (1972), "Punctuated Equilibria: an Alternative to Phyletic Gradualism", in T Schopf (ed.) *Models in Paleobiology*, Freeman Cooper, pp82-115.

Friedman D & Rust J (eds.)(1992), *The Double Auction Market: Institutions, Theories, and Evidence*, Addison-Wesley.

Harty K & Cheriton D (1996), "A Market Approach to Operating System Memory Allocation", in S Clearwater (ed.) *Market-Based Control: A Paradigm for Distributed Resource Allocation*, World Scientific, pp126-155.

Holland J H (ed.)(1975), *Adaptation in Natural and Artificial Systems*, University of Michigan Press.

Holland J H & Reitman J S (1978), "Cognitive Systems Based on Adaptive Algorithms", in D A Waterman & F Hayes-Roth (eds.) *Pattern-directed Inference Systems*, Academic Press.

Holland J H, Holyoak K J, Nisbett R E & Thagard P R (eds.)(1986), *Induction: Processes of Inference, Learning and Discovery*, MIT Press.

Holland J H & Miller J [1991], "Artificially Adaptive Agents in Economic Theory", *American Economic Review*, 81(2):365-370.

Kauffman S A & Johnsen S (1991), "Coevolution to the Edge of Chaos: Coupled Fitness Landscapes, Poised States, and Coevolutionary Avalanches", in C G Langton, J D Farmer, S Rasmussen & C Taylor (eds.) *Artificial Life II*, Addison-Wesley, pp325-370.

Lindgren K (1991), "Evolutionary Phenomena in Simple Dynamics", in C G Langton, J D Farmer, S Rasmussen & C Taylor (eds.) *Artificial Life II*, Addison-Wesley, pp295-312.

Marengo L & Tordjman H (1996), "Speculation, Heterogeneity and Learning: A Model of Exchange Rate Dynamics" *KYKLOS* 49(3):407-438.

Marimon R, McGrattan E & Sargent T (1990), "Money as a Medium of Exchange in an Economy with Artificially Intelligent Agents", *Economic Dynamics and Control* (14):329-373.

Miller M S & Drexler K (1988), "Markets and Computation: Agoric Open Syetems", in B Hubermann (ed.) *The Ecology of Computation*, Elsevier/North-Holland, pp133-176.

Miller M S, Krieger D, Hardy N, Hibbert C & Tribble E (1996), "An Automated Auction in ATM Network Bandwidth", in S Clearwater (ed.) *Market-Based Control: A Paradigm for Distributed Resource Allocation*, World Scientific, pp96-125.

Palmer R, Arthur W B, Holland J H, LeBaron B & Tayler P (1994), "Artificial Economic Life: A Simple Model of a Stockmarket", *Physica D* 75:264-274.

Potter M, De Jong K & Grefenstette J (1995), "A Coevolutionary Approach to Learning Sequential Decision Rules", in L J Eshelman (ed.) *Proceedings of the Sixth International Conference on Genetic Algorithms*, Morgan Kaufmann, pp366-372.

Seredynski F, Cichosz P & Klebus G (1995), "Learning Classifier Systems in Multi-Agent Environments", in *Proceedings of the First IEE/IEEE Conference on Genetic Algorithms in Engineering Systems: Innovations and Applications*, IEE, pp287-292.

Smith S F (1980), "A Learning System Based on Genetic Adaptive Algorithms", PhD dissertation, University of Pittsburgh.

Smith V (ed.)(1992), *Papers in Experimental Economics*, Cambridge Press.

Tomlinson A & Bull L (1998), "A Corporate Classifier System", in A E Eiben, T Baeck, M Schoenauer & H-P Schwefel (eds.) *Parallel Problem Solving from Nature - PPSN V*, Springer Verlag, pp550-559.

Watkins C (1989), "Learning from Delayed Rewards", PhD dissertation, University of Cambridge.

Wilson S W (1994), "ZCS: A Zeroth-level Classifier System", *Evolutionary Computation* 2(1):1-18.

Wilson S W (1995), "Classifier Fitness Based on Accuracy", *Evolutionary Computation* 3(2):149-176.