Evolution of Market Mechanism Through a Continuous Space of Auction-Types

Dave Cliff
Information Infrastructure Laboratory
HP Laboratories Bristol
HPL-2001-326
December 14th , 2001*

E-mail: dave_cliff@hp.com

genetic algorithm, market mechanism, zip traders, market design, mechanism design

A continuous space of auction mechanisms is explored via a genetic algorithm, with ZIP artificial trading agents operating in the evolved markets. The space of possible auction-types includes the Continuous Double Auction and also two purely one-sided mechanisms, yet *hybrids* of these auction types can be found to give the most desirable market dynamics.

* Internal Accession Date Only
Approved for External Publication

Evolution of market mechanism through a continuous space of auction-types

Dave Cliff
Hewlett-Packard Laboratories
Filton Road, Bristol BS34 8QZ, England, U.K.
dave_cliff@hp.com

Abstract: A continuous space of auction mechanisms is explored via a genetic algorithm, with ZIP artificial trading agents operating in the evolved markets. The space of possible auction-types includes the Continuous Double Auction and also two purely one-sided mechanisms, yet *hybrids* of these auction types can be found to give the most desirable market dynamics.

Abridged version of this paper submitted to CIFEr 2002.

I. INTRODUCTION

ZIP (Zero-Intelligence-Plus) artificial trading agents, introduced in 1997 [1], are software agents (or "robots") that use simple machine learning techniques to adapt to operating as buyers or sellers in open-outcry auction-market environments similar to those used in the experimental economics work of Smith (e.g. [2]). ZIP traders were originally developed as a solution to the pathological failures of Gode & Sunder's ZI (Zero-Intelligence) traders [3], but recent work by Das *et al.* at IBM [4] has shown that ZIP traders (unlike ZI traders) consistently out-perform human traders in human-against-robot experimental economics marketplaces.

The operation of ZIP traders has been successfully demonstrated in experimental versions of continuous double auction (CDA) markets similar to those found in the international markets for commodities, equities, capital, and derivatives; and in posted-offer auction markets similar to those seen in domestic high-street retail outlets [1,2]. In any such market, there are a number of parameters that govern the adaptation and trading processes of the ZIP traders. In the original formulation [1], the values of these parameters were set by hand, using "educated guesses". However, at CIFEr'98, the first results were presented from using a standard genetic algorithm (GA) to automatically optimise these parameter values [5], thereby eliminating the need for skilled human input in deciding the values of the parameters; more details of these results were subsequently given in [6].

In all previous work using artificial traders, ZIP or otherwise, the market mechanism (i.e., the type of auction the traders are interacting within) has been fixed in advance. Well-known market mechanisms from human economic affairs ~include: the English auction (where sellers stay silent and buyers quote increasing bid-prices), the Dutch Flower auction (where buyers stay silent and sellers quote decreasing offer-prices); the Vickery or second-price sealed-bid auction (where sealed bids are submitted by buyers, and the highest bidder is allowed to buy, but at the price of the *second-highest* bid -- game-theoretic analysis demonstrates that this curious mechanism encourages honesty and is robust to attack by dishonest means); and the CDA (where sellers announce decreasing offer prices while *simultaneously and asynchronously* the buyers announce increasing bid prices, with the sellers being free to accept any buyer's bid at any time and the buyers being free to accept any seller's offer at any time).

In this paper, the first results are presented from experiments where a GA optimises not only the parameter values for the trading agents, but also the style of market mechanism in which the traders operate. To do this, a space of possible market mechanisms has been created for evolutionary exploration. The space includes the CDA and also one-sided auctions similar (but not actually identical to) the English Auction (EA) and the Dutch Flower Auction (DFA); but significantly this space is *continuously variable*, allowing for any of an *infinite* number of peculiar hybrids of these auction types to be evolved, which have no known correlate in naturally occurring market mechanisms. While there is nothing to prevent the GA from settling on solutions that correspond to the known CDA auction type or the EA-like and DFA-like one-sided mechanisms, we repeatedly find that hybrid solutions are found to lead to the most desirable market dynamics. Although the hybrid market mechanisms could easily be implemented in online electronic marketplaces, they have not been designed by humans: rather they are the product of evolutionary search through a continuous space of possible auction-types. Thus, the results in this paper are the first demonstration that radically new market mechanisms for artificial traders may be designed by automatic means.

Section II gives an overview of our experimental methods, including a description of the continuously-variable space of auction types. In Section III we present our results, which are discussed in Section IV. Note that, in the descriptions that follow, we use $v=U[x,y]$ to denote a random real value v generated from a uniform distribution over the range $[x,y]$.

II. METHODS

A. Zero-Intelligence Plus (ZIP) Traders

ZIP traders are described fully in [1], which includes sample source-code in the *C* programming language. For the purposes of this paper a high-level description of the key parameters is sufficient. Each ZIP trader i is given a private (secret) limit-price, λ_i, which for a seller is the price below which it must not sell and for a buyer is the price above which it must not buy. If a ZIP trader completes a transaction at its λ_i price then it generates zero utility ("profit" for the sellers or "saving" for the buyers). For this reason, each ZIP trader i maintains a time-varying margin $\mu_i(t)$ and generates quote-prices $p_i(t)$ at time t according to $p_i(t)=\lambda_i\ (1+\mu_i(t))$ for sellers and $p_i(t)=\lambda_i\ (1-\mu_i(t))$ for buyers. The "aim" of traders is to maximise their utility over all trades, where utility is the difference between the accepted quote-price and the trader's λ_i value. Trader i is given an initial value $\mu_i(0)$ (i.e., $\mu_i(t)$ for $t=0$) which is subsequently adapted over time using a simple machine learning technique known as *the Widrow-Hoff rule* which is also used in back-propagation neural networks. This rule has a "learning rate" parameter β_i that governs the speed of convergence between trader i's quoted price $p_i(t)$ and the trader's idealised "target" price $\tau_i(t)$. When calculating $\tau_i(t)$, traders introduce a small random absolute perturbation generated from $U[0,c_a]$, and also a small random relative perturbation generated from $U[1-c_r,1]$ (buyers) or $U[1,1+c_r]$ (sellers) here c_a and c_r are system constants. To smooth over noise in the learning system, there is an additional "momentum" parameter γ_i for each trader (such momentum terms are also commonly used in back-propagation neural networks).

Thus, adaptation in each ZIP trader i has the following parameters: initial margin $\mu_i(0)$; learning rate β_i; and momentum term γ_i. In an entire market populated by ZIP traders, these three parameters are assigned to each trader from uniform random distributions each of which is defined via "min" and "delta" values in the following fashion: $\mu_i(0)=U(\mu_{min}, \mu_{min}+\mu_\Delta)$; $\beta_i=U(\beta_{min}, \beta_{min}+\beta_\Delta)$; and $\gamma_i=U(\gamma_{min}, \gamma_{min}+\gamma_\Delta)$. Hence, to initialise an entire ZIP-trader market it is necessary to specify values for the six market-initialisation parameters $\mu_{min}, \mu_\Delta, \beta_{min}, \beta_\Delta, \gamma_{min}$, and γ_Δ; and also for the two system constants c_a and c_r. And so it can be seen that any set of initialisation parameters for a ZIP-trader market exists within an eight-dimensional real space. Vectors in this 8-space can be considered as genotypes, and from an initial population of such genotypes it is possible to allow a GA to find new genotype vectors that best satisfy an appropriate evaluation function. This is exactly the process that was introduced at CIFEr'98 [5,6], and that is described further below.

When monitoring events in a real auction, as more precision is used to record the time of events, so the likelihood of any two events occurring at exactly the same time is diminished. For example, if two bid-quotes made at five minutes past nine are both recorded as occurring at 09:05, then they appear in the record as simultaneous; but a more accurate clock would have been able to reveal that the first bid was made at 09:05:01.64 and the second at 09:05:01.98. Even if two events occur absolutely at the same time, very often some random process (e.g. what direction the auctioneer is looking in) acts to break the simultaneity.

Thus, we may simulate real marketplaces (and implement electronic marketplaces) using techniques where each significant event always occurs at a unique time. We may choose to represent these by real high-precision times, or we may abstract away from precise time-keeping by dividing time (possibly irregularly) into discrete *slices*, numbered sequentially, where one significant event is known to occur in each slice.

In the ZIP-trader markets explored here, we use such a time-slicing approach. In each time-slice, the atomic "significant event" is one quote being issued by one trader and the other traders then responding either by ignoring the quote or by one of the traders accepting the quote. (NB in [4] a continuous-time formulation of the ZIP-trader algorithm was used).

In the markets described here and in [1,5,6], on each time-slice a ZIP trader i is chosen at random from those currently able to quote (i.e. those who hold appropriate stock or currency), and trader i's quote price $p_i(t)$ then becomes the "current quote" $q(t)$ for time t. Next, all traders j on the contraside (i.e. all buyers j if i is a seller, or all sellers j if i is a buyer) compare $q(t)$ to their own current quote price $p_j(t)$ and if the quotes cross (i.e. if $p_j(t)<=q(t)$ for sellers, or if $p_j(t)>=q(t)$ for buyers) then the trader j is able to accept the quote. If more than one trader is able to accept, one is chosen at random to make the transaction. If no traders are able to accept, the quote is regarded as "ignored". Once the trade is either accepted or ignored, the traders update their $\mu(t)$ values using the learning algorithm outlined above, and the current time-slice ends. This process repeats for each time-slice in a trading period, with occasional injections of fresh currency and stock, or redistribution of λ_i limit prices, until a maximum number of time-slices have completed.

B. Space of Possible Auctions

Now consider the case where we implement a ZIP-trader continuous double auction (CDA) market. In any one time-slice in a CDA either a buyer or a seller may quote, and in the definition of a CDA a quote is equally likely from each side. One way of implementing a CDA is, at the start of each time-slice, to generate a random binary variable to determine whether the quote will come from a buyer or a seller, and then to randomly choose one individual as the quoter from whichever side the binary value points to. Here, as in previous ZIP work [1][5][6] the random binary variable is always independently and identically distributed over all time-slices.

Let $Q=b$ denote the event that a buyer quotes on any one time-slice and let $Q=s$ denote the event that a seller quotes, then for the CDA we can write $Pr(Q=s)=0.5$ and note that because $Pr(Q=b)=1.0-Pr(Q=s)$ it is only necessary to specify $Pr(Q=s)$, which we will abbreviate to Q_s hereafter. Note additionally that in an English Auction (EA) we have $Q_s=0.0$, and in the Dutch Flower Auction (DFA) we have $Q_s=1.0$. Thus, there are at least three values of Q_s (0.0, 0.5, and 1.0) that correspond to three types of auction familiar from centuries of human economic affairs.

However, although the ZIP-trader case of $Q_s=0.5$ is indeed a good approximation to the CDA, the fact that any ZIP trader j will accept a quote whenever $q(t)$ and $p_j(t)$ cross means that the one-sided extreme cases $Q_s=0.0$ and $Q_s=1.0$ are not exact analogues of the EA and DFA. Nevertheless, consider the implications of considering values of Q_s of 0.0, 0.5, and 1.0 not as three distinct market mechanisms, but rather as three points on a *continuum*. How do we interpret, for example, $Q_s=0.1$? Certainly there is a straightforward implementation: on the average, for every nine quotes by buyers, there will be one quote from a seller. Yet the history of human economic affairs offers no examples (as far as I am aware) of such markets: why would anyone suggest such a bizarre way of operating, and who would go to the trouble of arbitrating (i.e., acting as an auctioneer for) such a mechanism? Nevertheless, there is no *a priori* reason to argue that the three known points on this Q_s continuum are the only loci of useful auction types. Maybe there are circumstances in which values such as $Q_s=0.712803$ (say) are preferred. Given the infinite nature of a real continuum, it seems appealing to use an automatic exploration process, such as the GA, to identify useful values of Q_s.

Thus, we add a ninth dimension to our search space, and the genotype in our GA is now the eight real values governing the ZIP-trader initialisation, plus a real value for Q_s.

C. The Genetic Algorithm

A simple genetic algorithm was used. In each experiment, we used a population of size 30 and allowed evolution to progress for 1000 generations. In each generation, all individuals were evaluated and assigned a fitness value; and the next generation's population was then generated via mutation and crossover on parents identified using rank-based selection. Elitism (where an unadulterated version of the fittest individual from each generation is copied into each successive generation) was also used.

The genome of each individual was simply a vector of nine real values. In each experiment, the initial random population was created by generating random values from $U[0,1]$ for each locus on each individual's genotype. Crossover points were between the real values, and crossover was governed by a Poisson random process with an average of between one and two crosses per reproduction event. Mutation was implemented by adding random values from $U[-m(g),+m(g)]$ where $m(g)$ is the mutation limit at generation g (starting the count at $g=0$). Mutation was applied to each locus in each genotype on each individual generated from a reproduction event, but the mutation limit $m(g)$ was gradually reduced via an exponential-decay annealing function of the form: $log_{10}(m(g))= - (log_{10} (m_s)-(g/(n_g-1)) log_{10}(m_s/m_e))$ where n_g is the number of generations (here $n_g= 1000$) and m_s is the "start" mutation limit (i.e., for $m(0)$) and m_e is the "end" mutation limit (i.e., for $m(n_g-1)$). In all the experiments reported here, $m_s=0.05$ and $m_e=0.0005$.

If ever mutation caused the value at a locus to fall outside [0.0,1.0] it was simply clipped to stay within that range. This clip-to-fit approach to dealing with out-of-range mutations has been shown [7] to bias evolution toward extreme values (i.e. the upper and lower bounds of the clipping), and so Q_s values of 0.0 or 1.0 are, if anything, more likely than values within those bounds. Moreover, initial and mutated genome values of μ_Δ, β_Δ, and γ_Δ were clipped where necessary to satisfy the constraints $(\mu_{min}+\mu_\Delta)<=1.0$, $(\beta_{min}+\beta_\Delta)<1.0$, & $(\gamma_{min}+\gamma_\Delta)<1.0$.

The fitness of genotypes was evaluated using the methods described in [5,6]. One *trial* of a particular genome was performed by initialising a ZIP-trader market from the genome, and then allowing the ZIP traders to operate within the market for a fixed number of trading periods, with allocations of stock and currency being replenished between trading periods. Each trading period ended either when no more trades are possible, or a maximum number of time-slices is reached.

During each trading period, Smith's α measure [2] of deviation of transaction prices from the theoretical market equilibrium price was monitored, and a front-weighted average was calculated across the trading periods in the trial. As the outcome of any one such trial is influenced by stochasticity in the system, the final fitness value for an individual was calculated as the arithmetic mean of 100 such trials. Note that as minimal deviation of transaction prices from the theoretical equilibrium price is desirable, lower scores are better: we are attempting here to *minimise* the fitness value.

Thus, in any one experiment, there are 30 individuals evaluated over 1000 generations where each evaluation involves calculating the mean of 100 trials, so a total of 3 million market trials would be executed in any one GA experiment (on a Hewlett-Packard Kayak XU800 workstation this would take approximately 5 hours). Nevertheless, the progress of each GA experiment is itself affected by stochasticity (e.g. the GA may become trapped on local minima) and so to generate reliable results each experiment was repeated 50 times (i.e., 150 million market trials, taking approximately 10.5 days). Results from 8 such 50-repeat experiments are shown below.

III. RESULTS

Figure 1 shows a supply and demand schedule for a marketplace with 11 buyers and 11 sellers, each empowered to buy/sell one unit of commodity: this market is referred to as Market 1. Figure 2 shows results from 50 repetitions of an experiment where the GA explores the 9-dimensional space: for each experiment, the fitness of the best (elite) member of the population is recorded. The results are clearly trimodal. Of the 50 repetitions, in 5 the elite ends up on fitness minima of about 3.2, while the other two elite fitness modes are on less-good minima of around 4.0 and 4.75. For comparison, Figure 3 shows the results of 50 repeats of the same experiment, where the value of Q_s was *not* evolved, being instead clamped at 0.5: i.e. the CDA value. The CDA mechanism is often applauded as an auction mechanism in which equilibration is rapid and stable, so we could expect the best fitness from using this market type. With the fixed CDA auction style, an average elite fitness of around 4.5 is settled on by the majority of experiments (48 repetitions) while a small minority (2 repetitions) settle on a less good mode of around

5.1. To ease the comparison, Figure 4 shows the data for the best elite fitness mode from Figure 2 (evolution of market mechanism) and the best elite fitness mode from Figure 3 (fixed CDA market mechanism) on the same graph.

Figure 5 shows a different supply and demand schedule, for the same number of traders (cf. Fig.1), referred to as Market 2. Figure 6 shows comparison of the mean scores from the best modes of 50 repetitions of an evolving-market (EM) experiment and 50 repetitions of a CDA fixed-market (FM) experiment (cf. Fig. 4) for Market 2. Similarly, Figure 7 shows a third supply and demand schedule, again for the same number of traders (cf. Figs 1 & 5), this being Market 3. Figure 8 shows comparison of the mean scores from the best modes of 50 repetitions of an EM experiment and 50 repetitions of a CDA-FM experiment for Market 3 (cf. Figs 4 & 6). It is clear from Figures 4, 6, & 8, that in each case the evolving-market experiments are significantly better than the CDA fixed-market experiments. Figures 9, 10, & 11 show the evolutionary trajectory of the mean evolved value of Q_s in each of these three experiments.

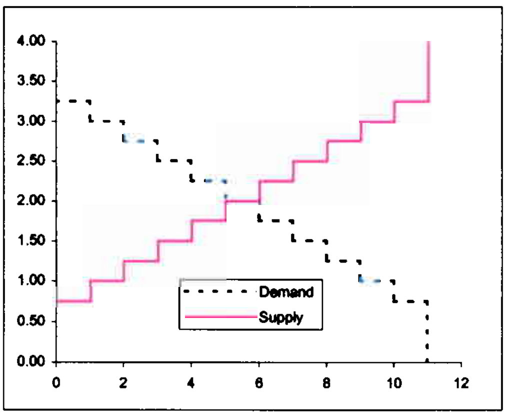

Figure 1: Supply and demand schedules for Market 1.

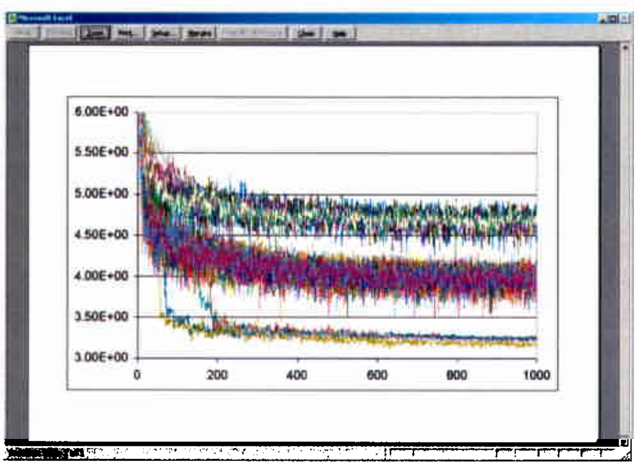

Figure 2: Elite fitness values from 50 repetitions of the 1000-generation evolving-market (EM) experiment operating with Market 1. Lower values are better solutions (less deviation from equilibrium). Results are trimodal, with five of the repetitions (10%) settling to values around 3.2.

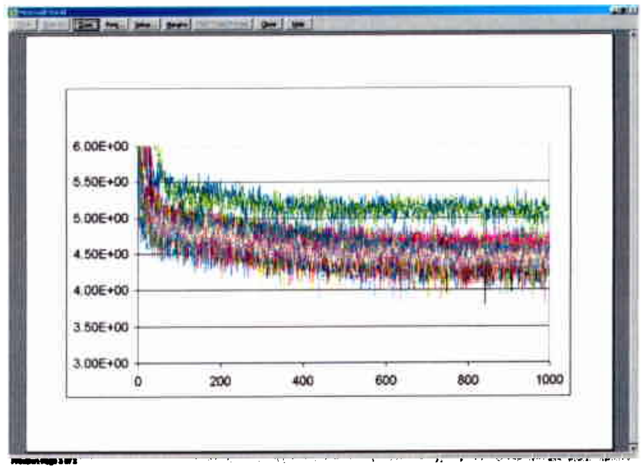

Figure 3: Elite fitness values from 50 repetitions of a 1000-generation experiment operating with Market 1, but with a fixed-market (FM) CDA mechanism with $Q_s=0.5$: bimodal results, with 96% of the repetitions settling to fitness values around 4.5 and the remaining 4% at around 5.2.

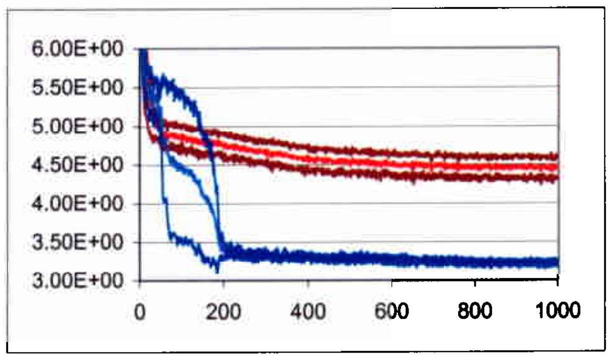

Figure 4: Comparison of mean elite fitnesses in the best solution modes from Figs. 2 and 3. For both sets of experiments, data is plotted for mean elite fitness, plus and minus one standard deviation (s.d.). EM fitnesses settle to a mean of approx. 3.2 with a s.d. of approx 0.02 (n=5); FM fitnesses settle to a mean of just under 4.5 with a s.d. of approx. 0.15 (n=48).

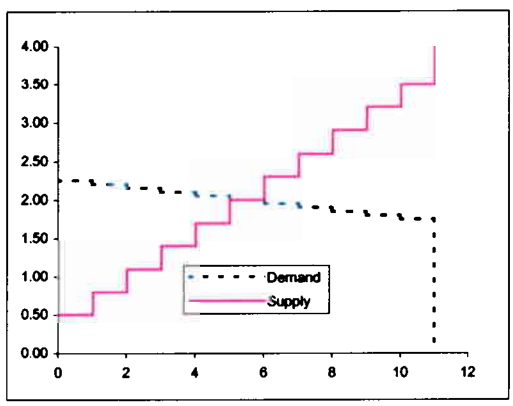

Figure 5: Supply and demand schedules for Market 2.

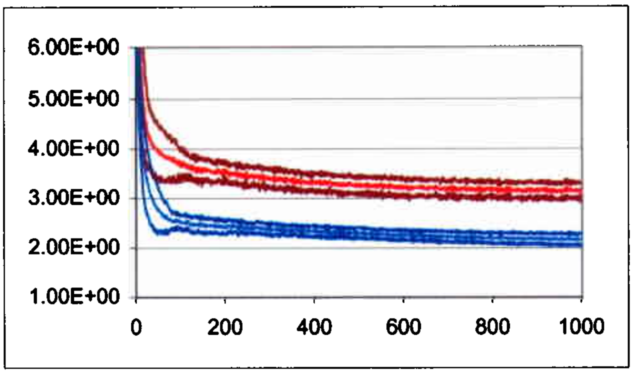

Figure 6: Average elite fitnesses from 50 EM and 50 FM(Q_s=*0.5*) experiments for Market 2; data is plotted for mean fitness, plus and minus one s.d.. Best EM fitness mode settles to a mean of just over 2.15 with a s.d. of approx 0.11 (n=45); FM values settle to a mean of around 3.14 with a s.d. of approx. 0.16 (n=50).

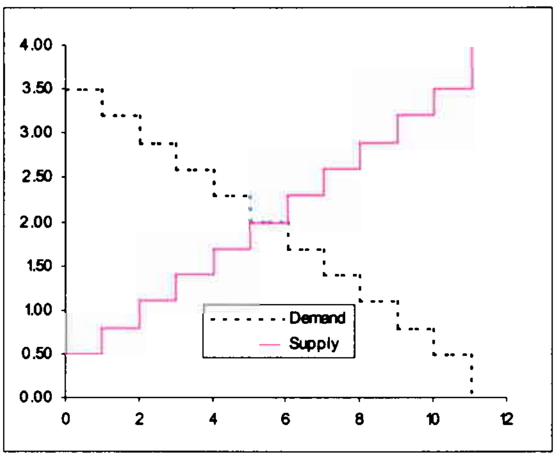

Figure 7: Supply and demand schedules for Market 3.

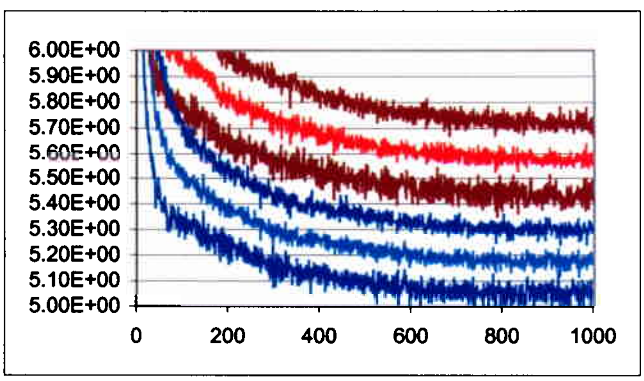

Figure 8: Comparison of mean fitnesses in the best solution modes from 50 EM and 50 FM (Q_s=*0.5*) experiments with Market 3; data is plotted for mean elite fitness, plus and minus one s.d.. EM elite fitnesses settle to a mean of just over 5.18 with a s.d. of approx 0.12 (n=50); FM elite fitnesses settle to a mean of around 5.55 with a s.d. of approx. 0.15 (n=50).

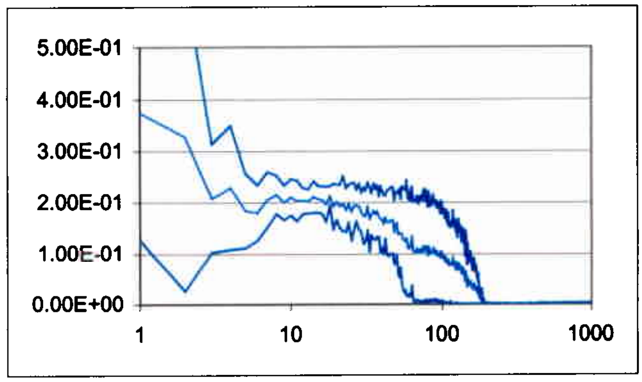

Figure 9: Evolutionary trajectory of mean (plus and minus one standard deviation) value of Q_s in the Market 1 EM experiments settling on the best fitness mode (n=5): mean settles to Q_s values around 0.0001 (the average value of the mean over generations 900 to 1000 is 0.0000894).

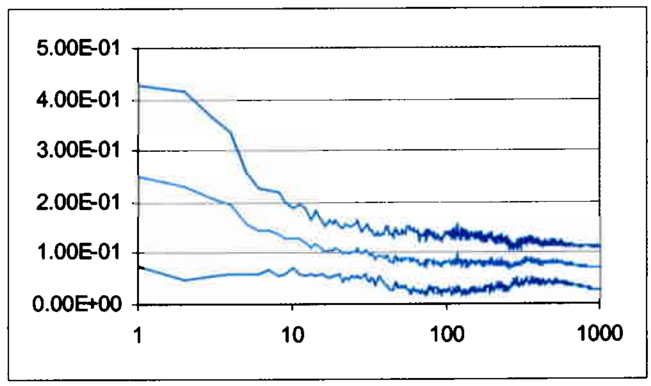

Figure 10: Evolutionary trajectory of mean (plus and minus one s.d.) value of Q_s in the Market 2 EM experiments settling on the best fitness mode (n=45): mean value of Q_s is approx 0.07.

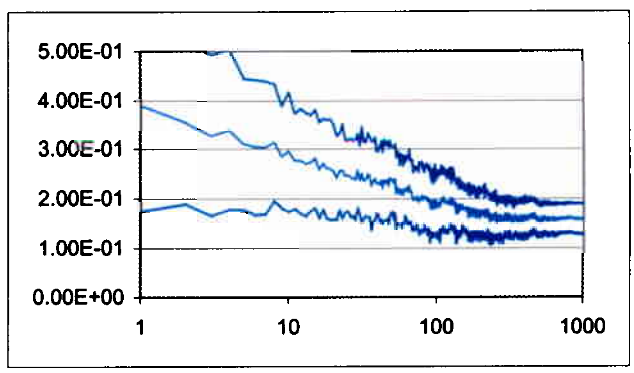

Figure 11: Evolutionary trajectory of mean (plus and minus one s.d.) value of Q_s in the Market 3 EM experiments settling on the best fitness mode (n=50): mean value of Q_s is a little under 0.16.

Notably, in only one of the three experiments does the value of Q_s settle at a clear value such as 0.0, 0.5, or 1.0: the Market 1 results are sufficiently close to zero that we can consider the outcome in Market 1 to be $Q_s=0.0$. However, in the two other markets, the Q_s values settled on do not correspond to values known from any existing market mechanisms.

Finally, Figure 12 contrasts the average β_{min} values (i.e., the lower bound on the traders' learning rates) in the two sets of Market 3 experiments. As can be seen, the values of β_{min} appropriate for the fixed CDA market are significantly higher than the values found to give most fitness in the evolving-market case. Such variations in final evolved ZIP parameter values between the evolving-market and fixed-market cases are common. Significant differences in evolved parameter values can also occur between the outcomes of fixed-market experiments where evolution adapts the ZIP parameters to the specific market supply and demand schedules used.

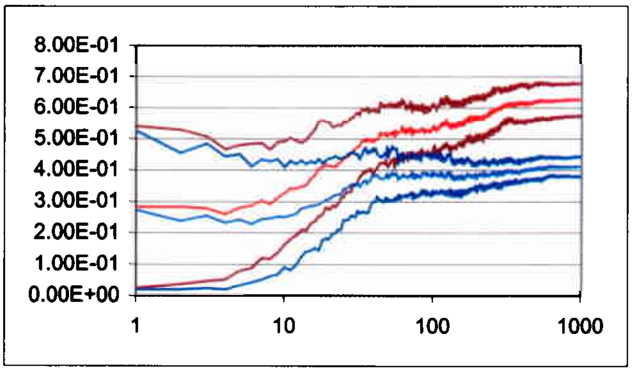

Figure 12: Evolutionary trajectory of mean (plus and minus one s.d.) values of β_{min} in the Market 3 experiments. FM ($Q_s=0.5$) values settle on a mean of around 0.62 while EM values settle on a mean of approx 0.4.

IV. DISCUSSION

For Market 1, the value of Q_s is so close to zero that we can consider the evolved mechanism as a one-sided auction with rules similar to those of the English auction, except that there are multiple sellers simultaneously active, each with different reservation prices, and any of which can "intervene" to accept a bid at any time (rather than waiting until only one bidder remains).

For Markets 2 and 3, let us approximate the final mean Q_s value in Market 2 (Fig. 10) as 0.0714 and the final mean Q_s for Market 3 (Fig. 11) as 0.1666. Taking reciprocals of these two values, we find that, despite their being an equal number of buyers and sellers in both of these markets, the evolved mechanism for Market 2 dictates that on average only one quote in 14 should come from a seller; and the evolved mechanism for Market 3 requires that on average only one quote in 6 should come from a seller.

As far as we are aware, no human-designed markets follow such rules, so these two mechanisms can reasonably be described as unlike any market mechanism previously devised by humans. Yet the fitness data for these two evolved market mechanisms clearly demonstrate that they give less deviation of transaction prices from the equilibrium price than the fixed CDA design does.

However, the Q_s trajectory for Market 1 (shown in Figure 9) raises the possibility that the final evolved Q_s values for Markets 2 and 3 are misleading. That is: perhaps in Markets 2 and 3 a value of $Q_s=0.0$ would give even better fitness values; and maybe evolution never reaches this extreme value because, once the value of Q_s falls below a certain threshold, the selective pressure for further reductions is sufficiently diminished (or masked by noise) that the value of Q_s simply drifts on a random walk in the sub-threshold zone until the end of the experiment. If the nonzero evolved values of Q_s in Markets 2 and 3 are to be taken seriously, the possibility that better fitness results might be obtained by fixing Q_s at zero needs to be explored.

Thus, two further sets of 50 fixed-market (FM) experiments were conducted: one for Market 2 and one for Market 3. In each, the value of Q_s was fixed at zero for the duration of the experiment, and all other ZIP-trader parameters were subject to evolution as before. Individual fitness results for Market 2 are shown in Fig.13 and for Market 3 in Fig.14. The mean values from these experiments are shown in Figs 15 and 16.

The mean fitnesses in the best solution modes from these two sets of FM ($Q_s=0.0$) experiments are very close to the mean values of the best modes in the corresponding evolving-market (EM) experiments. In Market 2 (Fig.15) the mean final best-mode fitness value in the FM ($Q_s=0.0$) experiments is approximately one standard deviation lower (i.e., better) than in the corresponding EM market (Fig.6, lower trace), and hence we conclude that for Market 2 the EM experiments failed to find a solution that yields statistically lower results (i.e., better fitness) than the FM case for the fixed one-sided mechanism of $Q_s=0.0$. That is, the final evolved value of $Q_s=\sim0.0714$ is not significantly better than if $Q_s=0.0$ had been chosen *a priori.*

However, in Market 3, there is a more intriguing result. Here, the mean final best-mode fitness value in the FM ($Q_s=0.0$) experiments is approximately one standard deviation *higher* than in the corresponding EM market (Fig.8, lower trace). Given that we have a large number of samples (i.e. around 50 repetitions in each case), it is very plausible that this is an indication that the evolved market is significantly better than the fixed one-sided case. To rigorously test this possibility, we used the Wilcoxon form of the Wilcoxon-Mann-Whitney Test [8, p.128ff.], which is preferred to Student's *t* test as it is nonparametric (i.e., makes no assumption about the underlying sample distributions). The null hypothesis H_0 is that the final mean best-mode fitnesses in the EM and FM($Q_s=0.0$) experiments are drawn from the same population, and the alternative hypothesis H_1 is that the FM values are stochastically larger than the EM values (i.e., that the evolved market gave *lower* fitness values, and hence *better* market dynamics). The directional nature of H_1 requires a one-tailed test, and the large N of 98 (=48+50) makes the normal approximation appropriate. The z-value for the samples is 2.8746, which is significant at the 0.005 level, and hence we safely reject H_0 and accept H_1 (see Appendix).

That is to say, for Market 3, the evolved value Q_s=~0.1666, which results in a hybrid CDA/one-sided market where on average one quote in six comes from a seller, does indeed give statistically better values than either the CDA or the pure one-sided $Q_s=0.0$ market mechanism. Thus, it can be concluded that the evolved hybrid market mechanism is best suited (among those considered) to giving the desired market dynamics (i.e., minimal deviation between transaction-price and equilibrium price) in Market 3.

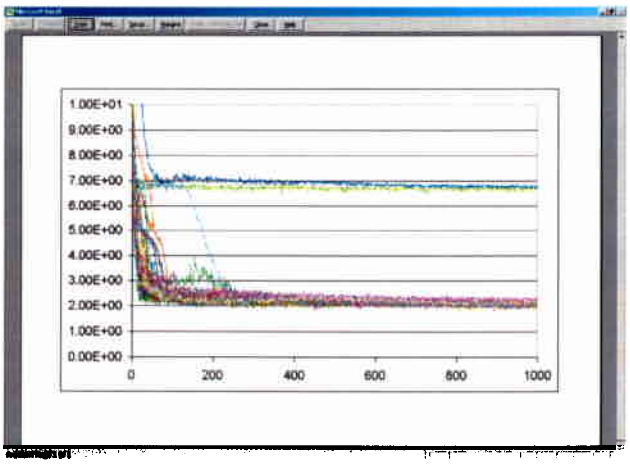

Figure 13: Elite fitness values from 50 repetitions of the 1000-generation FM ($Q_s=0.0$) experiment operating with Market 2. Results are bimodal, with three of the repetitions (6%) settling to values around 6.8, and the remaining 47 repetitions settling on a mode of approximately 2.1.

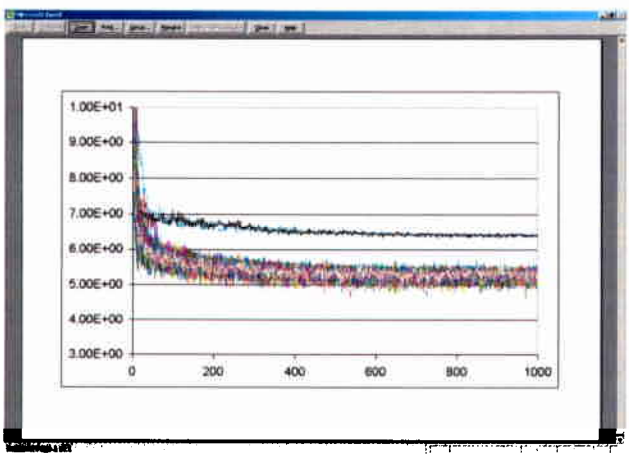

Figure 14: Elite fitness values from 50 repetitions of the 1000-generation FM ($Q_s=0.0$) experiment operating with Market 3. Results are bimodal, with two of the repetitions (4%) settling to values around 6.4 and the remaining 48 (96%) settling on a mode of approximately 5.2.

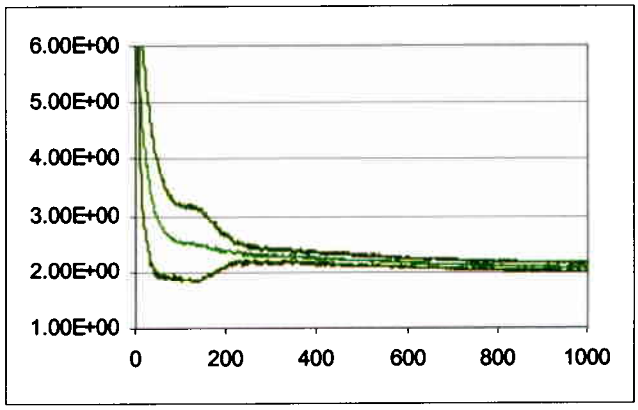

Figure 15: Mean elite fitnesses in the best solution mode from 50 FM ($Q_s=0.0$) experiments with Market 2 (cf. Fig. 13); data is plotted for mean fitness, plus and minus one s.d.. FM fitnesses settle to a mean of around 2.08 with a s.d. of approx. 0.08 (n=47).

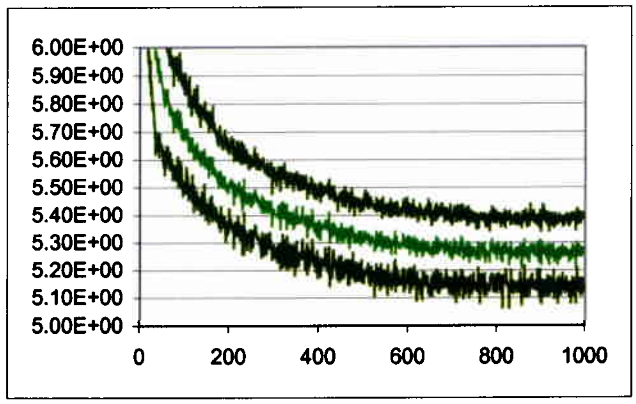

Figure 16: Mean elite fitnesses in the best solution mode from 50 FM ($Q_s=0.0$) experiments with Market 3 (cf. Fig. 14); data is plotted for mean fitness, plus and minus one s.d.. FM fitnesses settle to a mean of around 5.27 with a s.d. of approx. 0.12 (n=48).

V. CONCLUSION

In this paper we have demonstrated the use of an evolutionary search through an infinite space of possible market designs that includes the CDA of $Q_s=0.5$ and also the two pure one-sided solutions of $Q_s=0.0$ and $Q_s=1.0$. A new "hybrid" market mechanism was found to give the most desirable market dynamics in one of the three experiments. While such evolved market mechanisms are unlike any human-designed mechanism, they could nevertheless readily be implemented as online electronic marketplaces.

Future work will be directed toward formulating more sophisticated "genetic encodings" for richer spaces of possible market mechanism, and in developing a firm understanding of why and under what conditions the evolved hybrid markets give better market dynamics than existing human-designed mechanisms. Variations in the evolved market mechanisms resulting from altering the fitness function will also be explored. In particular, attempting to maximise buyer utility (or seller utility) is likely to result in the evolution of mechanism designs significantly different to those presented here.

Acknowledgements

Thanks to Andrew Byde for a discussion that clarified the nature of the $Q_s=0.0$ and $Q_s=1.0$ one-sided auction mechanisms, and to Rycharde Hawkes for his co-operation with my intensive use of an ad-hoc compute-farm.

References

[1] D. Cliff, "Minimal-intelligence agents for bargaining behaviours in market environments" Technical Report HPL-97-91, Hewlett-Packard Labs. (http://www.hpl.hp.com/techreports/97/HPL-97-91.html)

[2] V. Smith, "Experimental study of competitive market behavior" *Journal of Political Economy* **70**:111-137, 1962.

[3] D. Gode & S. Sunder, "Allocative efficiency of markets with zero-intelligence traders" *Journal of Political Economy* **101**:119-137, 1993.

[4] R. Das, J. Hanson, J. Kephart, & G. Tesauro ,"Agent-human interactions in the continuous double auction" *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI-01)*, 2001. (http://www.research.ibm.com/infoecon/researchpapers.html)

[5] D. Cliff, "Genetic optimization of adaptive trading agents for double-auction markets" in *Proceedings of Computational Intelligence in Financial Engineering (CIFEr) 1998*. IEEE/IAFE/Informs (preprint proceedings), pp.252-258, 1998.

[6] D. Cliff, "Evolutionary optimization of parameter sets for adaptive software-agent traders in continuous double-auction markets", Technical Report HPL-2001-99, Hewlett-Packard Laboratories. (http://www.hpl.hp.com/techreports/2001/HPL-2001-99.html)

[7] S. Bullock, "Are artificial mutation biases unnatural?" in: Floreano, D, Nicoud, J-D & Mondada, F (eds) *Advances in Artificial Life - Fifth European Conference (ECAL99)*, pp. 64-73 Springer-Verlag. 1999. (http://www.comp.leeds.ac.uk/seth/papers/ecal99_mutation.ps.gz)

[8] S. Siegel & N. J. Castellan, *Nonparametric Statistics for the Behavioural Sciences*. McGraw Hill. 1988.

Appendix

The table below shows data for the Wilcoxon version of the Wilcoxon-Mann-Whitney Test used on the Market 3 EM and FM($Q_s=0.0$) results. The m=48 final best-mode fitness scores from the FM($Q_s=0.0$) experiment (Fig.16) and the n=50 final best-mode fitness scores from the EM experiment were grouped together, with the EM values marked as Type 1 and the FM values marked as Type 2. Fitness values were then assigned a rank-order based on their position following sorting into ascending order. There were no tied ranks.

Summing the rank values for Type 1 (EM) gave a value W_1 and summing the rank values for Type 2 (FM) gave a value W_2. Reading the columns from left to right, the table shows: the rank value; the fitness value; the source type of the fitness value; the cumulative value of W_1; and the cumulative value of W_2. The final rank sums are W_1=2071 and W_2=2780.

Let $N=m+n$; then using $z=(W_2+0.5-m(N+1)/2)/\mathrm{SQRT}(mn(N+1)/12)$ gives z=2.874656, while the 0.005 one-tailed significance level is 2.576.

			n=50	m=48
			W_1	W_2
Rank	Fitness	Type	0	0
1	4.89669	1	1	0
2	4.93447	1	3	0
3	4.94673	1	6	0
4	4.96113	2	6	4
5	4.96787	2	6	9
6	5.00727	1	12	9
7	5.01772	1	19	9
8	5.02246	2	19	17
9	5.02557	1	28	17
10	5.04571	1	38	17
11	5.05407	1	49	17
12	5.06159	1	61	17
13	5.07337	1	74	17
14	5.08056	1	88	17
15	5.08379	1	103	17
16	5.10063	1	119	17
17	5.10252	1	136	17
18	5.10713	2	136	35
19	5.10831	1	155	35
20	5.11039	1	175	35
21	5.11301	1	196	35
22	5.11844	2	196	57
23	5.12229	2	196	80
24	5.13012	1	220	80
25	5.13117	2	220	105
26	5.13120	2	220	131
27	5.13856	1	247	131
28	5.14950	2	247	159
29	5.15147	1	276	159
30	5.15318	1	306	159
31	5.15435	1	337	159
32	5.15529	2	337	191
33	5.16116	1	370	191
34	5.16402	1	404	191
35	5.17092	2	404	226
36	5.17579	1	440	226
37	5.19243	2	440	263
38	5.19408	2	440	301
39	5.19648	1	479	301
40	5.19665	1	519	301
41	5.20168	2	519	342
42	5.21755	2	519	384
43	5.21976	1	562	384
44	5.22043	1	606	384
45	5.22983	1	651	384
46	5.23013	2	651	430
47	5.23031	2	651	477
48	5.23191	1	699	477

49	5.23263	2	699	526
50	5.23558	2	699	576
51	5.23821	1	750	576
52	5.23971	2	750	628
53	5.23972	2	750	681
54	5.24590	2	750	735
55	5.24766	1	805	735
56	5.25612	2	805	791
57	5.26601	1	862	791
58	5.26639	1	920	791
59	5.27560	1	979	791
60	5.27719	2	979	851
61	5.28281	2	979	912
62	5.28322	2	979	974
63	5.28602	2	979	1037
64	5.28603	2	979	1101
65	5.28636	1	1044	1101
66	5.28672	2	1044	1167
67	5.29027	2	1044	1234
68	5.29339	1	1112	1234
69	5.29341	1	1181	1234
70	5.29580	2	1181	1304
71	5.30351	2	1181	1375
72	5.31535	1	1253	1375
73	5.31539	1	1326	1375
74	5.31819	1	1400	1375
75	5.32491	1	1475	1375
76	5.32529	2	1475	1451
77	5.32746	2	1475	1528
78	5.32755	2	1475	1606
79	5.32992	1	1554	1606
80	5.33327	1	1634	1606
81	5.33441	2	1634	1687
82	5.34111	2	1634	1769
83	5.34298	2	1634	1852
84	5.34990	1	1718	1852
85	5.36183	1	1803	1852
86	5.36402	2	1803	1938
87	5.36493	1	1890	1938
88	5.37816	1	1978	1938
89	5.39061	2	1978	2027
90	5.39237	2	1978	2117
91	5.40863	2	1978	2208
92	5.43686	2	1978	2300
93	5.43876	1	2071	2300
94	5.44567	2	2071	2394
95	5.46467	2	2071	2489
96	5.47440	2	2071	2585
97	5.48235	2	2071	2682
98	5.49331	2	2071	2780