CO-EVOLUTION OF TRADING STRATEGIES IN AN ARTIFICIAL COMPUTATIONAL STOCK MARKET

Sergio Margarita[1] and Andrea Beltratti[2]

[1] Istituto di Matematica Finanziaria, Università di Torino, Via Assarotti 3, 10122 Torino, Italia; tel. +39 11 546805; fax +39 11 544004
[2] Istituto di Economia Politica G. Prato, Università di Torino, Via della Cittadella 10/E, 10122 Torino, Italia; tel. +39 11 540900; fax +39 11 541497

Abstract

We describe a financial market populated by different species of agents who trade in order to increase their wealth. The dynamics of both populations is governed by simple behavioral and learning rules. We use artificial neural networks to model the decision processes of these agents who learn and act, and interpret the input-output mappings coming out of the learning process as time-invariant decision rules to describe the resulting evolutionary process in terms of co-evolution of strategies. The idea of interpreting the aggregate outcome of the market as the final results of an evolving process of agents following simple rules is related with the Artificial Life literature. The results show that the aggregate market outcome emerging from repeated interactions of these simple agents is indeed complicated, and difficult to interpret without an a priori knowledge of the behavioral rules of the agents. Our simulations show that such strategies co-evolve over time in a manner that is related to the target assigned to each agent, that is the maximization of wealth.

1. Introduction

Financial markets are populated by agents who trade in order to maximize their profits; to a certain extent the interaction has a zero-sum game aspect, since often the agents who are better at evaluating the assets are the ones who attract the larger proportion of wealth. In practice agents take decisions under conditions of imperfect information and imperfect knowledge, so that the structure of the decision-making problem is likely to be bounded rational. Economists have responded to the challenge of understanding the dynamics of financial markets with models that make many restrictive hypotheses about the behavior of traders: they are often assumed to have perfect knowledge of the models of the economy and act with rational expectations.

We believe that such modelling strategy is useful mainly for situations of stable environments that are hardly found in actual financial markets (see section 2), and advocate the use of an Artificial Life-based approach to the study of financial markets. Instead of explaining a complex aggregate outcome with a complex set of equations that are unlikely to describe the way agents behave in practice, we consider agents that behave according to simple rules taking into account computational limits to the elaboration of information, and analyze the emerging structure. In our model agents are heterogeneous also with respect to their institutional role in the market; there is a dealer who has the purpose of buying and selling shares at prices that she herself decided at the beginning of the trading session under conditions of incomplete information about the market demand function. We believe that the model is relevant for economists and for decision-makers, since it illustrates a process where imperfect knowledge about the general conditions of the market is of central concern to a decision-maker that has to act every period on the basis of the available information and of the current imperfect knowledge of the system.

The structure of the paper is the following: section 2 describes the importance of an Artificial Life approach for analyzing financial markets, section 3 describes the economic structure of the market we are considering, while section 4 discusses some issues that are relevant to the effort of modelling agents as artificial neural networks.

Section 5 and 6 describe more precisely the behavior and the learning process of both the traders and the dealer. Section 7 presents the results and section 8 concludes.

2. Why an AL-based approach for financial markets

Economists who study financial markets by means of analytical models have to make some very restrictive assumptions on the behavior and on the computational ability of the agents in order to obtain closed form solutions. The standard assumptions are that agents buy and sell assets according to a well defined demand function and that all the parameters of the system are well known to the decision-makers. Common knowledge of the true model of the economy is known how the hypothesis of rational expectations, that requires an incredible computational ability on the part of all the agents. In order to solve the resulting models economists often use the fiction of a representative agent.

This methodology is certainly useful for classes of problems where the environment is relatively stable and the parameters of the various equations are constant and have been estimated by the agents (Lucas 1986). There are many cases in which learning about a changing environment with computational limitations is the main part of the problem of the decision-maker (Winter 1986). For example the financial market is a continuously changing system, where agents act on the basis of incomplete information. Such agents have to perform on-line learning: they cannot wait for the end of the history of the economy before making decisions, since in many cases decisions are effective only if taken in few seconds. Accordingly we believe that financial markets are particularly ill-suited to the analytic methodology that may be valid for stable environments.

In the search for an alternative approach we first note that the structure of the model (heterogeneous agents competing for the appropriation of a resource by adopting various possible strategies) recalls co-evolutionary models of the biological literature, where, according to Koza (1992) "... the environment of a given species includes all the other biological species that contemporaneously occupy the physical environment and which are simultaneously trying to survive". We then note that in such biological models the description of the behavior of the agents is usually very simple; they process information in a local way and behave according to rules of thumb derived on the basis of their computational limits. As Langton (1992) and others have shown in what is now known as the Artificial-Life approach, the interactions of such simple agents may give rise to very complex aggregate outcomes. We follow this methodology and choose to model agents behaving according to simple rules that may be implemented with algorithms requiring a modest computational ability. Our agents try to maximize their wealth over time by learning appropriate mappings between their information sets and their targets, without obeying to any predefined rule. They do not know the state of the overall system, and take decisions on the basis of a small subset of the available information. Such simple rules give rise to a complex aggregate behavior that we try to interpret on the basis of simple strategies defined in terms of observable variables (see section 7).

Note that the idea of co-evolution has already found important applications in the analysis of social systems: for example Miller (1989), Hillis (1990), Kephart, Hogg and Huberman (1990). This paper continues on the one hand the line of research of Arthur (1991), Holland and Miller (1991) and further explores the consequences of heterogeneity in financial markets in a model that analyzes the outcomes of the interactions between a group of traders (along the lines of Beltratti and Margarita 1992) and a dealer. For a general discussion of the applicability of Artificial Life to financial markets see Margarita (1992).

3. The on-screen stock market: the rules of the game

We are interested in studying the dynamics of an on-screen stock market, where all the traders can buy and sell shares from a dealer. The last behaves as a price-maker and decides at the beginning of each trading session the price she is willing to offer for buying shares from the traders (the bid price) and the price she is willing to ask for selling shares to the traders (the ask price). The difference between the ask and the bid is the spread, that gives the dealer a compensation for the job of operating the market. Traders are better off if they increase their wealth, which they can do by buying low and selling high. The dealer is better off if she accumulates more wealth, which she can do by increasing the spread. The interests of the two agents are therefore opposite: the larger the spread, the better off the dealer, the worse off the traders. There is however a limit to how far the spread can go, since when the spread is too high traders do not join the market and stop trading. We let the number of traders who join the market in any given day depend negatively on the spread decided by the dealer. In such a way the last faces a trade-off between earning more profit for each transaction and earning less unitary profit on a larger number of transactions. The process that limits the ability of the dealer to charge a very large spread could be certainly modelled as the result of a process of competition among different dealers. However we consider only one dealer to simplify the simulation structure of our model, that would become very involved with more than one dealer, since each trader would need to compare the prices offered and asked by all the dealers before trading with the one offering the best conditions.

4. Agents as artificial neural networks

To model the behavior of agents we use artificial neural networks (ANN). The neurons of the various layers are connected with a set of weights, that are modified in a recursive way in order to adapt the output to an externally set target for each period of time. Such modifications are achieved by means of a back-propagation algorithm (Rumelhart, McClelland 1986 and White 1991).

The basic advantage of modelling economic agents as ANNs is to allow for learning over time and action on the basis of (i) the available information (ii) decision rules which are not imposed externally. The main drawbacks are (i) that it is not always possible to devise targets related to outputs of the network, in order to implement supervised learning and (ii) that it is often difficult to interpret in economic terms the decision rules adopted by the network.

As to (i), one can well say that building appropriate targets for supervised learning is the more involving task when using ANNs as a description of economic agents. In some cases (like for the traders in our model) economic theory suggests a certain target, and at the same time helps to interpret the outputs of a network in terms of decisions taken by a rational agent. In other cases (like for the dealer in our model, who should know the demand function to take the wealth-maximizing decision about prices) economic theory suggests a target that is well defined in theoretical terms, but that in practice would require an incredible amount of information and computational ability on the part of the agent. In section 6 we will point out the need to integrate the standard learning by back-propagation with elements that are typical of a reward-and-punishment methodology, in order to describe a learning process with the merits of easy computability and of being close to a reasonable rule of thumb that may well be followed in practice by agents with a limited computational ability. (ii) instead refers to the possibility of interpreting the input-output mapping of a given network in terms of strategies, that is in terms of time-invariant description of the reaction of a network to a given set of signals coming from the environment. In fact the structure of a network implies that inputs and outputs are real numbers; this is indeed one of the main advantages of using neural networks instead of classifier systems that may only consider discrete alternatives and are therefore less general from a computational point of view. This lower generality however also implies a better interpretability of the input-output mapping in terms of decision rules. The approximation we use in this paper, described in detail in the following sections, consists of arbitrarily interpreting inputs and outputs in terms of variation and/or discrete choices in order to give a classification of the possible decisions. The discretization of the inputs and outputs space may be used to attain an approximate description of the decision rules of the traders without losing the benefit of modelling agents who can learn and act in a continuous space. Agents learn continuously within a given strategy, and may sometime switch from one strategy to another; there are times when an incremental learning has no effect on the strategy, and others when a small change in the learning process brings about a sudden change in strategy.

5. The traders

(A) Behavior: trader i enters the early morning of day t with a given amount of money $M_{i,t}$, a stock of shares $S_{i,t}$ and an information set $I_{i,t}$. The latter includes P_{Bt}, D_t, P_{Bt-1}, D_{t-1}, $|Q_{it-1}|$, $sgn(Q_{it-1})$, where P_{Bt} is the bid price decided by the dealer at the beginning of time t, at which it is possible to sell any amount of shares to the dealer, D_t is the spread decided by dealer at time t, that drives a wedge between bid and ask prices, since $P_{At} = P_{Bt} + D_t$, Q_{it-1} is the amount of shares transacted with the dealer in the last period, $sgn(Q_{it-1})$ is the sign of the operation, equal to +1 for a buy and -1 for a sell (Figure 1). Each day traders can buy (sell) shares against money in a stock market, at a price P_{At} (P_{Bt}) which is decided by the dealer. Agents decide the amount of shares they want to trade, with the purpose of increasing their wealth. The exact mechanism that induces traders to make wealth increasing decisions is described in (B) below.

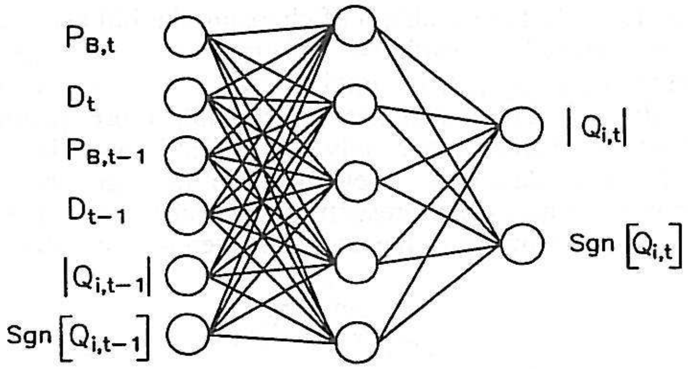

Figure 1. Structure of the traders

(B) Learning: how do traders learn over time? Economic theory, and common sense, suggests that traders who maximize wealth try to buy (sell) before the price goes up (down), in order to have the option to sell (buy) at a higher (lower) price in the future. We follow this simple but rational rule and teach the traders to buy (sell) any time that the information set signals a future increase (decrease) in market prices. We define $P_t=(P_{At} + P_{Bt})/2$, and compare the decision taken by traders at t-1 with the change P_t-P_{t-1}; if the change was positive we tell the trader she should have bought more shares at the beginning of day t-1. We compare the buy/sell decision of the network at t-1 with the *right* one implied by the change in the future price; if the actual and the optimal action coincide there is no error and no learning about this specific output, otherwise we signal the network an error that gives rise to learning.
As to learning about the size of the transaction (the amount of shares that the network decided to buy or sell), we compare the actual number of shares that was traded with the number that should have been traded in order to maximize wealth, taking into account the constraint that each agent can buy shares only if enough money is available, and cannot sell shares that are not already owned. This optimal number can be determined by considering the sign of the change in price from one period to the following:
(i) If this change was positive, the optimal action in t-1 was a buy and the target is equal to the maximum trade that would have been allowed by the budget constraint in t-1, that is M_{t-1}/P_{Bt-1};

(ii) If instead the change was negative, the optimal action in t-1 was a sale, and the target is equal to S_{t-1}, that is the number of shares that were held at the beginning of day t-1, and that should have been completely sold before a drop in price.

(C) Interpreting the decision rules of traders: in order to interpret the decisions taken by the traders in terms of time-invariant decision rules we consider the decision to buy or sell as a function of the following inputs:

change in bid from t-1 to t
change in spread from t-1 to t
decision (to buy or sell) taken in t-1

This would yield a large number of possible decision rules, that are however restricted to 4 when the decision of the dealer about the bid and the spread is announced at the beginning of day t. For example if the dealer announced a bid that is larger than the one of period t-1 and a spread that is lower than the one prevailing in t-1 the trader has 4 possible combinations that relate the buy or sell decision at time t as a function of the action at t-1 (buy/buy, buy/sell, sell/buy, sell/sell). In this way we can characterize 4 *conditional* strategies that show persistence or reversal of behavior from one day to the other.

6. The dealer

(A) Behavior: the dealer has the problem of choosing the bid and the ask price, again with the goal of increasing her wealth. Her information set is composed of Q_{Bt-1}, Q_{St-1}, P_{Bt-1}, D_{t-1}, where Q_{Bt-1} (Q_{St-1}) is the total amount of shares that were bought (sold) by the dealer in day t-1. The dealer has considerable more information than each single trader about the market, since only the dealer knows the total demand and supply at the end of each day, and can evaluate the imbalance between demand and supply. The output of the network representing the dealer is P_{Bt} and D_t (Figure 2). The next section clarifies the structure of the decision taken by the dealer.

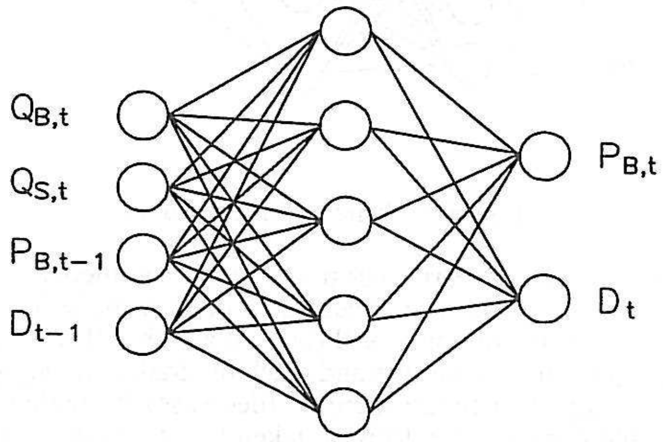

Figure 2. Structure of the dealer

(B) Learning: learning of the dealer is more involved than learning of traders. The consequences in terms of wealth of decisions about the spread and the price depend on the reaction of the environment, that is of all the traders. If the dealer decides to increase the bid price and the spread and her wealth decreases, it is not clear that a decrease in prices would have increased wealth. The problem here is that the agent does not know the relationship between actions and wealth. There is a double problem of learning the consequences of one's actions and of learning which actions actually increase wealth.

We consider 4 possible vectors of strategies $Z_{i,t}$, whose first element is the change in the bid price and the second element is the change in the spread: $Z_{1,t}=(-1,-1)$, $Z_{2,t}=(-1,+1)$, $Z_{3,t}=(+1,-1)$, $Z_{4,t}=(+1,+1)$. There is a time-dependent associated vector ($p_{1,t}$, $p_{2,t}$, $p_{3,t}$, $p_{4,t}$) representing the probabilities of adoption of each strategy. At each time

t the network representing the dealer decides a bid price and a spread. These values will be larger or lower than the ones prevailing in time t-1, so that the decision can be classified in one of the four possible strategies. Trading takes place at the resulting prices and wealth of the dealer goes up or down:
(i) If the change in wealth is positive, there is an indication that the strategy that was selected was working in the right direction, so the dealer maintains the same strategy and we increase the probability associated with this strategy. The targets for price and spread at time t, aimed at giving a reinforcement, are defined in the following way:

$$\begin{array}{llll} \text{Price target} & = & P_{Bt} + \alpha(P_{max} - P_{Bt}) & \text{if } P_{Bt} - P_{Bt-1} > 0 \\ & = & P_{Bt} - \alpha(P_{Bt} - P_{min}) & \text{if } P_{Bt} - P_{Bt-1} < 0 \\ \\ \text{Spread target} & = & D_t + \beta(D_{max} - D_t) & \text{if } D_t - D_{t-1} > 0 \\ & = & D_t - \beta(D_t - D_{min}) & \text{if } D_t - D_{t-1} < 0 \end{array}$$

(ii) If the change in wealth is negative instead, the followed strategy was wrong and the dealer changes strategy by randomly selecting a strategy among the other 3 possible according to the vector of probabilities at time t; then we decrease the probability of the strategy that has been abandoned by the dealer.
If the new strategy corresponds to the vector $Z_{j,t}$, we define targets for price and spread in the following way, in order to give a punishment:

$$\begin{array}{llll} \text{Price target} & = & P_{Bt-1} + \alpha(P_{max} - P_{Bt-1}) & \text{if } Z_{j,t}(1)=+1 \\ & = & P_{Bt-1} - \alpha(P_{Bt-1} - P_{min}) & \text{if } Z_{j,t}(1)=-1 \\ \\ \text{Spread target} & = & D_{t-1} + \beta(D_{max} - D_{t-1}) & \text{if } Z_{j,t}(2)=+1 \\ & = & D_{t-1} - \beta(D_{t-1} - D_{min}) & \text{if } Z_{j,t}(2)=-1 \end{array}$$

where $Z_{j,t}(1)$ ($Z_{j,t}(2)$) represents the first (second) component of vector Z.
The idea of the scheme is that the neural network acts by herself as long as the change in wealth is positive; the probability of the strategy that is providing this change in wealth is increasing at the expense of the probabilities of the other 3, and the target is reinforcing the action that was actually taken: an increase in the price (spread) that brought about an increase in wealth is compared with a larger increase, to encourage the network to take a stronger decision in the same direction in the future. When the strategy fails the dealer decides to change her strategy by choosing a new strategy with a probability proportional to its past success. So this mechanism corrects the decision rule, by decreasing the probability that the strategy is chosen and by giving a target compatible with the strategy that is suggested as an alternative for the future.
(C) Interpreting the decision rules of the dealer: we adopt the following description of the mapping between inputs and outputs based on four binary variables:
(i) the sign of the change in the bid price from t-1 to t
(ii) the sign of the change in the spread from t-1 to t
(iii) the sign of the quantity traded in t-1
(iv) the sign of the quantity traded in t.
We define a decision rule as a mapping between the first three elements and the last.

7. Results

We report our results in terms of figures 3-8, describing the results of a simulation. Figure 3 reports the dynamics of the bid and ask prices, showing that there is much volatility in the decisions of the dealer as far as the prices are concerned. There is much less volatility in the spread, the vertical difference between the two lines (note however that spread is constrained between 0.01 and 0.06). Figure 4 reports the strategy chosen by the dealer at each point of time, together with the evolution of her wealth. The dealer seems to use different couples of strategies for different periods. From 0 to 50 she alternates between strategies 2 and 4, from 50 to 100 she uses 1 and 3, then switches with decision to 4 until 150, and then mostly uses 2 and 3, a couple

that had never appeared before. There seems to be some relationship between the chosen strategy and the dynamics of wealth, with more experimentation of new strategies when wealth is more stable. Towards the end of the sample for example there is some experimentation with strategy 1.
Figure 5 reports the proportions of traders using the various possible strategies, that is continue to sell, continue to buy, switch from buying to selling, switch from selling to buying.

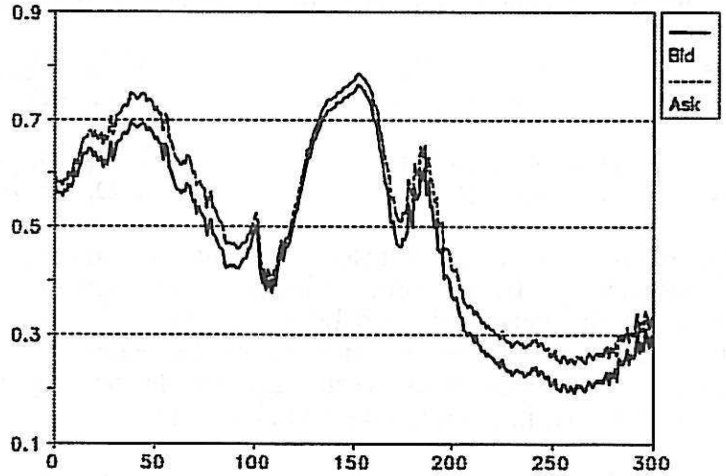

Figure 3. Dynamics of bid and ask prices

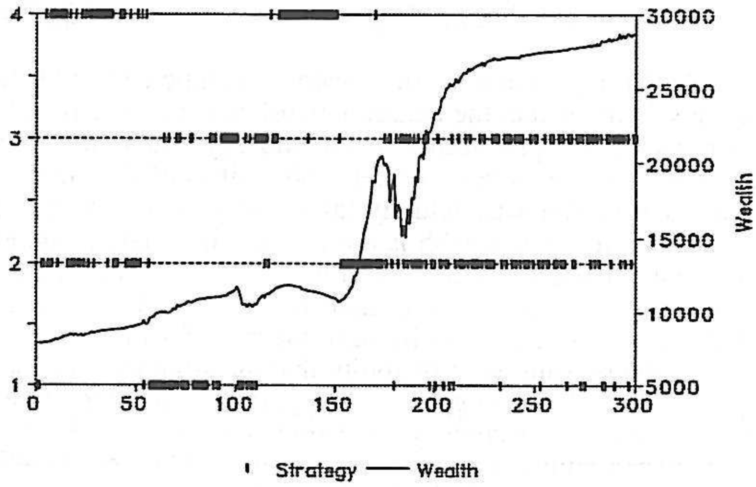

Figure 4. Strategies and wealth of the dealer

There is a clear predominance of the strategies that suggest to keep doing the same thing from one period to the other. From 50 to 120 almost everybody keeps buying, while from 170 to 240 everybody keeps selling. What is the relationship between this pattern and the strategies chosen by the dealer? Figure 6 reports the bid price and the difference between the proportion of traders who buy (among those who trade in every given day) and the proportion of traders who sell (among those who trade in every given day). There is a strong co-movement between the two series; most of the traders buy when the price is rising since their extrapolation of the past trend in prices suggests them that there is persistency in the tendency of prices, so that when the price starts rising there are good chances that the price will keep on rising.

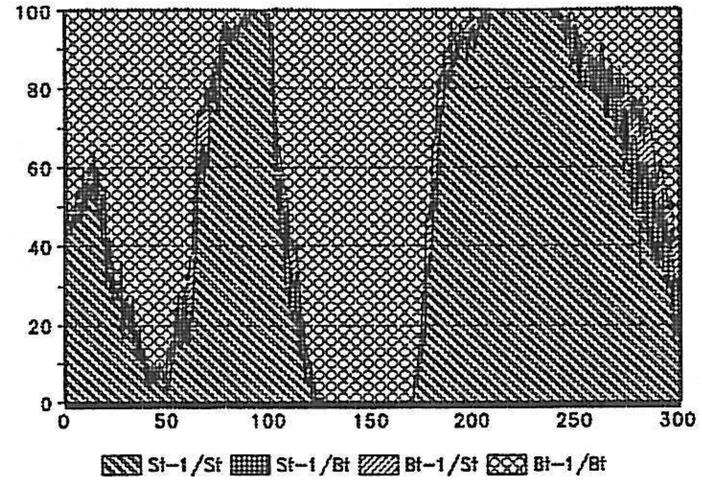

Figure 5. Evolution of traders' strategy

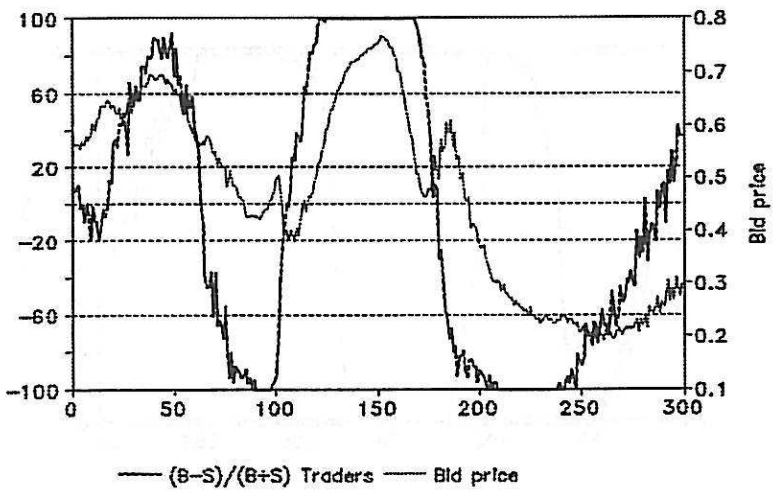

Figure 6. Proportion of buying traders and bid price

There are therefore clear relationships in the decisions of the traders and those of the dealer. Figures 7 and 8 comment more on this point, by showing the co-evolution of the strategies of the dealer and of the population of traders. In order to simplify the figure we consider for the dealer separately price strategy (figure 7) and spread strategy (figure 8) versus the proportion of selling traders. In both figures, value -1 corresponds to a decrease while value 1 denotes an increase.

We can describe the mutual evolution of strategy of dealer and traders in the following way (for the sake of simplicity, we only focus on the dealer's price strategy, without commenting on the spread strategy):

a) Let us suppose that dealer starts following a successful strategy corresponding to an increase of price;

b) Her wealth increases and she does not change strategy;

c) Market price tends to rise;

d) In response to dealer's strategy, by the effect of learning rules, buyers overcome sellers in the population of traders;

e) Dealer is forced to increase sales and reduce purchases;

f) Wealth of the dealer, computed as $W_{d,t} = M_{d,t} + S_{d,t} P_t$, is exposed to two opposite forces: a positive one due to the increase in price, and a negative one due to the reduction of the number of owned shares;

g) When all the traders become buyers, the latter becomes stronger and the dealer has to face a decrease in her wealth;

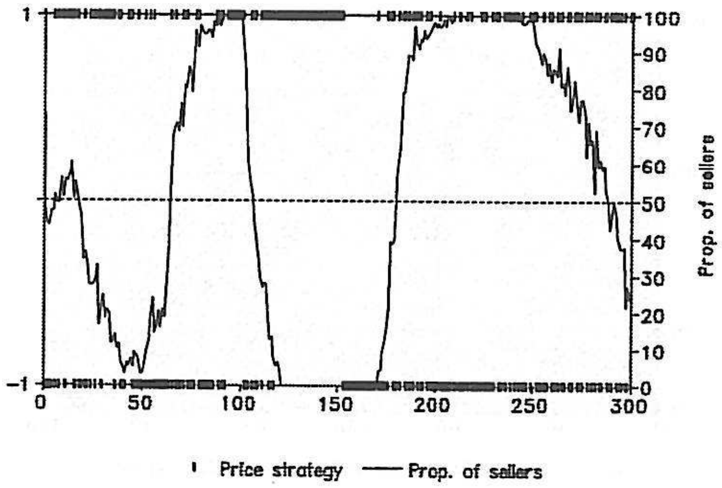

Figure 7. Co-evolution of traders' and dealer's strategies
(Dealer's price strategy)

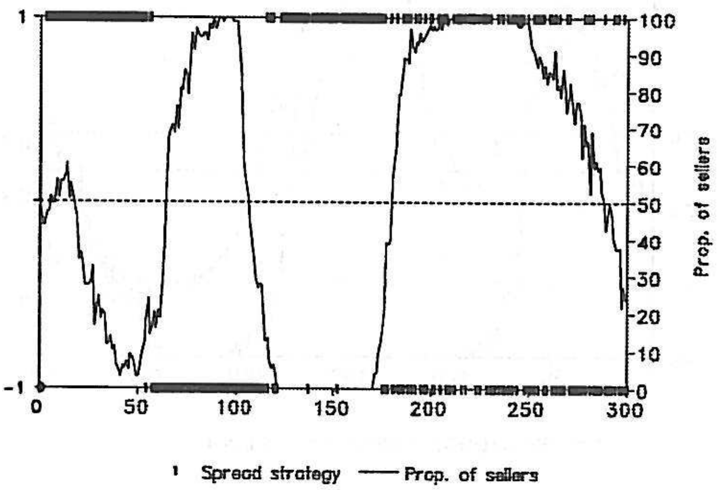

Figure 8. Co-evolution of traders' and dealer's strategies
(Dealer's spread strategy)

h) She reacts by changing her strategy and begin to reduce bid price;
i) Reacting to this new strategy, sellers reappear on the market and help to confirm the decreasing trend;
j) Wealth of dealer increases and she tends to maintain this new strategy;
k) Sellers become dominant in the market;
l) After few periods, a phenomenon similar to the one described in f), but with opposite sign occurs which generates a new reversal in the strategy of the dealer;
m) Long-run dynamics of the system is generated by iteration of these steps.
Similar observations can be done about spread strategy. Interactions between price strategy and spread strategy have yet to be explored.
Two facts contribute to smooth the dynamics of the system:
(i) The swaps that have been described refer to learning strategy, that is to the one used for building the targets of dealer. So they have a delayed effect on real strategy;
(ii) Traders often exhibit differences in individual behavior (learning speed, reactivity to inputs) which can generate some spikes in the evolution of the dealer's wealth. In some periods in fact, dealer does not use a single strategy but a couple of strategies, swapping frequently from one to the other.

8. Conclusions

Actual stock markets present complex dynamics; this Artificial Life-based approach tries to understand them by means of the synthesis of single agents' behavior rather than the analysis of the complex global outcome of the system. One aspect that deserves more attention is the influence of learning on the building of behavioral rules: it is very important to distinguish the part of real emergent behavior from the one implicitly induced by learning rules.
The results show that the aggregate market outcome emerging from repeated interactions of these simple agents is indeed complicated, and difficult to interpret without an a priori knowledge of the behavioral rules of the agents. Such interpretation is made more easy by our attempt of describing the structure of the various networks in terms of strategies defined on a discretized space of actions. Our simulations show that such strategies co-evolve over time in a manner that is related to the target assigned to each agent, that is the maximization of wealth. Future directions of research will focus on multi-population models, with many dealers against many traders, and models with different goals for dealers: not only wealth maximization but also compensation of imbalance between global demand and supply in the market, in accordance to real behavior of professional dealers on some stock markets.

References

Arthur W.B., 1991, Designing economic agents that act like human agents: a behavioral approach to bounded rationality, American Economic Review, vol. 81, n.2, pp. 353-359

Beltratti A. and S. Margarita, 1992, Evolution of trading strategies among heterogeneous artificial economic agents, in J.A. Meyer, H.L. Roitblat, S.W. Wilson (eds.), From animals to animats 2, Proceedings of the second international conference on simulation of adaptive behavior, Cambridge, Mass., MIT Press (forthcoming)

Hillis W.D., 1992, Co-evolving parasites improve simulated evolution as an optimization procedure, in C.G. Langton et al., Artificial Life II, Redwood City, Addison-Wesley

Holland J.H. and J.H. Miller, 1991, Artificial adaptive agents in economic theory, American Economic Review, vol. 81, n.2, pp. 365-370

Kephart J.O., T. Hogg and B.A. Huberman, 1990, Collective behavior of predictive agents, Physica D 42, pp. 48-65

Koza J.R., 1992, Genetic evolution and co-evolution of computer programs, in C.G. Langton et al., Artificial Life II, Redwood City, Addison-Wesley

Langton C.G., 1992, Life at the edge of chaos, in Langton et al., 1992, Artificial Life II, Redwood City, Addison-Wesley

Langton C.G., C. Taylor, J.D. Farmer and S. Rasmussen, 1992, Artificial Life II, Redwood City, Addison-Wesley

Lucas R.E., 1986, Adaptive behavior and economic theory, Journal of Business, vol. 59, pp. 401-426

Margarita S., 1992, Interacting neural networks: an artificial life approach to stock markets, in I. Aleksander, J. Taylor, Artificial neural networks 2, Amsterdam, North-Holland

Miller J.H., 1989, The co-evolution of automata in the repeated prisoner's dilemma, Santa Fe Institute Report 89-003

Rumelhart D.E. and J.L. McClelland, 1986, Parallel distributed processing, Explorations in the microstructure of cognition, Cambridge, Mass., MIT Press

White A., 1991, Learning in artificial neural networks: a statistical perspective, Neural Computation, vol. 1, pp. 425-464

Winter, S.G., 1986, Comments on Arrow and Lucas, Journal of Business, vol. 59, pp. 427-434