Detecting State Transitions in a Stock Market with Many Agents

Eric Van Horenbeeck
CNTS - Language Technology Group
University of Antwerp
ericvh@uia.ua.ac.be

Abstract

The attitude of traders in a stock market when they have to decide on what price to trade, is not unlike the aggregated behavior of individual fish or birds: random swarm phases are interchanged with polarized and deterministic schooling phases. The changing from one state to another is often induced by small fluctuations accumulating to a macroscopic transition. The noise-induced shift from the deterministic phase to the random phase is treated as a loss of memory. Schooling is interpreted as a learning experience where the system 'remembers' where it has been. *Random Harvester* is an trading agent that detects the phase transitions between swarming and schooling and uses this knowledge to his advantage. Preliminary results are presented.

Keywords: stock market, trading agent, swarm, schooling, state transition

1 Introduction

How do traders forecast the path that a particular stock will take? According to a widely accepted view it is done either with technical rules or with fundamental analysis. Technical rules are based on the assumption that prices move in a predictable direction in a pattern derived from analysis of historical data. Fundamental forecasting assumes that prices reflect the discounted worth of the stocks' future returns. In this communication a novel way to decide when to buy or when to sell is presented based on the analysis of a stock market as a complex system. Although such a system is inherently difficult to predict, the agent-based trading model perceives the point where the ascent of a particular stock stops and changes into a descent. Augmented with the conventional business rules of the stock trade (transaction costs, expiration dates, etc.) the agent makes a profitable trade in 85% of the cases during a simulation with historical data of 100 stocks over a two year period. The first four paragraphs present the assumptions used in the development of the agent. Parts 5 to 7 outline some important technical characteristics. Finally paragraph 8 shows a few results of simulated trading with real (historical) data.

2 The Stock Market Does Not Follow Random Walks.

A large body of research exists on how to describe adequately the behavior of stock market prices. The Random Walk theory instructs us that stock prices are fundamentally independent from each other: any knowledge about a price at a certain moment does not add to the knowledge of a price on another moment. It is (was?) the dominant doctrine among academics and professionals (Lo and MacKinlay, 1999). A subsequent conclusion based on the underlying probability theory is that all price series have the same distribution. It is of little importance at what specific instant the series starts: any moment its as good as the next. The best bet for tomorrow's stock price is its value today. *'If a stock price was going to go up 10$ tomorrow, it would go up 10$ today'*. A variant of the Random Walk is the Efficient Market Hypothesis. This theory asserts that the ideal investor has a perfect foresight and uses only information on asset prices that is consistent with his rational expectations. Prices are assumed to be rationally related to economic realities. (De Grauwe *et al.*, 1993).

These premises are continuously debated by researchers who publish findings that reject or severely impair the Random Walk Hypothesis. Empirical observations suggest that the succession of daily, weekly and monthly distributions progressively converge to a Gaussian distribution. The longer the interval in the series, the more they seem to move up and down smoothly as is illustrated by Figure 1. But to use this pattern to improve the accuracy of a forecast of the Down Jones is an illusion of arithmetic. Large variations in stock prices happen with sufficient frequency to raise doubts about existing models which fail to account for non-Gaussian statistics.

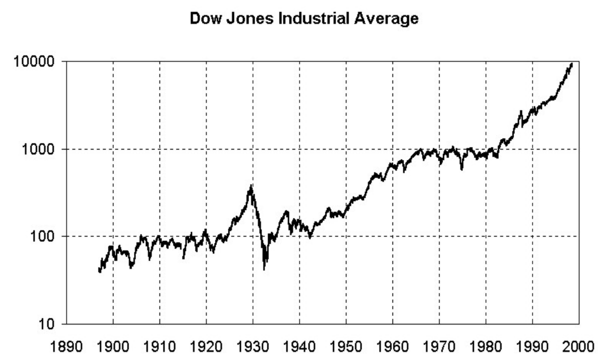

Figure 1: Long intervals create an illusion of predictability

Edgar Peters (1991) is one of several authors who therefore characterize the stock markets as a nonlinear dynamic system with:

1. Long-term feedback effects and trends;
2. Erratic markets under certain conditions, and at certain times;
3. A time series of returns with fractal structure that will look the same and will have similar statistical characteristics at smaller increments of time;
4. Sensitive dependence on initial conditions and thus less reliable forecasts, the further out in time.

Mandelbrot observed that price variations of many market indices over different, but relatively short time intervals could be described by a stable Lévy distribution. The Lévy distribution has substantially more weight for large events than the Gaussian distribution where large events beyond 4σ are unlikely (Mandelbrot, 1966). For example, the data for the S&P 500 index is reasonably fitted by a truncated Lévy distribution with $\alpha \simeq 1.4$ over a time scale which ranges from a minute to a day with convergence to a Gaussian at approximately one month (Mantegna and Stanley, 1995). In recent work Mandelbrot distances himself even further from any hope of finding distribution normality in stock prices (Mandelbrot, 1998).

The large variations may be due to a crowd effect, where agents imitate each other's behavior. In a departure from the traditional view with only one type of economic participant, Bak *et al.* (1996) at Santa Fé devised a model with two types of agents trading stock. Type (1) agents are 'noise' traders whose current volatility depends on recent changes in the market and whose choice of price to buy or sell imitates choices of others; type (2) are 'rational' agents optimizing their own utility functions. This function U has the form:

$$U = \nu\, min[A, B] + (1-\nu)(\rho A + (1-\rho)B)$$

In the equation every individual has an overall risk adversity utility function which is a mix of extreme risk aversion ($\nu = 1$) and risk neutral ($\nu = 0$). The probability distribution of a transaction A (buy) or B (sell) is expressed by ρ. In the model there is only one type of stock and each agent can own at most one share. In spite of these simplifications the model is capable of exhibiting a statistical pattern of price variations consistent with that observed empirically. If the model is loaded with only rational traders who chose prices according to the utility function U then the market evolves into a state where all N/2 agents who own stock have a selling price p_s and all N/2 agents who do not own any shares have a buying prices p_b. Buyers and sellers are soon locked in non-overlapping regions of price with a dead zone in between. Thus, with only one stock for sale the rational agents will not trade in equilibrium but hold their portfolios indefinitely. When noise traders take over the toy stock exchange and act in a purely random manner, the model corresponds to the universality class of the reaction diffusion process A + B $\rightarrow$ 0, i.e. the market evaporates due to stocks at a price 0. However when the noise agents can imitate rational traders and each other on a trade floor the model escapes this universality class and exhibits behavior characteristic of real markets.

3 The Stock Market is a Self-Organizing System

The communication of information towards and between economic agents is postulated to be a major driver of trading behavior (French and Roll, 1986). Again there is abundant literature that offers evidence on the widely and frequent deviation between actual prices and the rational valuation of an asset (Schiller, 1984). In the early studies this 'abnormalities' are often explained away as errors, fads or bubbles. The individual agent gets the blame and the Efficient Market Hypothesis goes scot-free. The Santa Fé model suggests that the arrival of external information is not necessary to provoke market volatility. How could that be? If there is an unclear connection between events in the economy and their effects on the stock exchange, predicting what will happen in the future with a reasonable degree of accuracy is difficult if not impossible. Yet there is considerable regularity of behavior. Systems seem to be capable of organizing themself in spite of the unpredictable interactions between agents in and out of that system. Where the system will be at any point in time is hard to tell, but often bounds can be set around the areas in which it will move.

Enters the ant. Ants have become the favorite little animals of many a researcher in the AI field. The dead ant test of Deneubourg (1990) shows that the coordinated behavior of piling dead ants in a single pile results from a simple local rule controlling the behavior of a single ant. The piling behavior is accounted for without postulating communication between the ants. Well known is also the model of Kirman (Kirman in (Arthur *et al.*, 1997)) who translates ant feeding behavior to the financial markets with agents acting either on the fundamentals of a stock or on a technical indicator.

4 Shifting between order and disorder

So far for the self-organizing capacity of a system. Remains the fact that stock markets do show pockets of regularity in a sea of randomness. When at sea, look for fish.

Fish schools are highly organized. They do not need a leader or external stimuli to form a polarized structure. The members show a parallel arrangement with a uniform distance between individuals and synchronized motion. In a paper on the self-organizing features of fish schools Hiro-Sato Niwa (1994) makes a distinction between fish behaving as a swarm and fish behaving as a school. The correlation between the velocity of a particular fish and those of other fishes in a school is not strong. Speed and heading are not closely related to those of any other single fish. The strong correlations are observed between the velocity of the individual and average velocity of the entire school. The individual then adjusts its own velocity to match an average. Thus in a sense the entire school is the leader and an individual is a follower. Field observations found that when a fish accidentally swims outward from a school in the state of Brownian movement (random walk) its nearest neighbor companions draw toward it and equalize their velocities. If this initial fluctuation is sufficiently large the order locally appearing in a school grows rapidly to a high degree of parallel orientation throughout the whole school. The author approximates this behavior with a non-linear Langevin equation near the instability point. There is a critical relation between the arrayal coupling constant J and the strength of the noise of individual movements ϵ, such that in the swarming phase, where $J < 5\beta\epsilon$ the center of the school hardly moves, whereas in the schooling phase with $J > 5\beta\epsilon$, the fish form a tighter group with a rectilinear movement. β shows the non-linearity of the system, $\beta^{-1/2}$ is the steady swimming speed of fish. As the strength of the interaction becomes large at $J\beta^{-1/2} = 5\epsilon/\beta^{-1/2}$, the schooling structure is self-organizing. In the equation $\epsilon/\beta^{-1/2}$ stands for the magnitude of random movement in a school and $J\beta^{-1/2}$ denotes the mean strength of the influence one individual by the other individuals as a group. The interplay between the fluctuations and non-linearities of the system leads to unexpected modifications of the macroscopic behavior, from polarized schooling to random swarming and vice versa. The transition from random behavior to schooling is important but equally interesting is the fact that the order can be broken at any moment.

5 Trading as swarming and schooling

Back to the stock market. Trading occurs at prices formed in the course of a trading day.

Prices are a collection of discrete entities with a certain distribution. The collection of trading prices moves as a group in a field. Sometimes the movement could be qualified as random (swarming) at other moments the movement is definitely polarized and synchronous (schooling). Researchers noticed that most intraday fluctuations are self-generating rather than explained by the arrival of sensible information (Harris, 1986). Our own research shows that the path of a group of trading prices for a particular asset can be described by its centroid: they move about their center just as fish do. Those intraday prices are at the heart of the stock trade, all other time scales are statistical reductions that loose a lot of information in exchange for a simplified image of reality.

Fitting the pieces together:

1. Individual traders do not explore the complete pricing space. The mix of rational utility maximalizers and imitators (noise traders) leads to the clustering necessary for a market to be dynamic and liquid. This is achieved without formal communication between the members.

2. The clusters show self-organizing features: without much hints a group polarizes and changes from Brownian indetermination (swarm phase) into a deterministic rally (schooling phase) and back again. Information arrival can sometimes trigger such a phase transition, but it doesn't have to.

3. The alternation of swarming and schooling phases plotted on a timescale exhibits all the properties of a fractal series. Fractals imply a connection between the different points of the series. Random numbers are not connected. In a fractal each point is correlated with an earlier point, if that were not the case then a random point tracker would fill the space or plane it is visiting.

4. The history of a stock price is a vector field in a plane of all possible prices. A trader analyses the price path hoping to find forms of stability that leads him to the detection of the underlying attractor of the vector field. Both the technical and the fundamental analyst presume the existence of a periodic pattern of the limit circle type: given a set of conditions gathered from a previous period, the future course can be predicted (Fama and French, 1992). Forecasts of this kind are hazardous. The main reason why it often goes wrong is the instability of the path.

6 Random Harvester, a Trading Agent

The *Random Harvester* approach to stock trading conjectures the existence of a system where patches of non-linear functioning (random walk - swarm phase) are alternating with stretches of well structured action (schooling phase). Both the duration and the order of succession of each pattern is not known in advance. The solution to predicting the behavior of a trading community at $t+1$ from knowledge of the system in t is not to try to model the full path, but to look for state transition points. Stock prices either have a random signature in the swarming phase or sustain a reciprocal relation for some time. The algorithm should find out what is the case. To do this the agent starts with an empty memory and is fed with the timestamp - trade price pairs from a particular asset. The incoming prices make a collection that grows as long as the new augmented collection passes the normality test. In this sense the set of normally distributed data represents the memory of the system. It is used to make a calculated guess on what to expect for the next period. When new data violate this expectation the system forgets its past experiences and reports a state transition. In brief three pieces of information are needed to construct the picture: the way the data are distributed, a measure of the normality of the local data set and a slope function to indicate whether the trading path is going upwards or downwards.

When trading is behaving deterministic and the slope is > 0 a rectangular wave is created with an arbitrary value of +1. When the slope is < 0 the value is -1. When no state transition is observed the wave value is not changed and stays at the preceding value of +1 or -1. A buying signal is generated when a state transition from random swarming to polarized behavior is detected. A sell signal is given when the schooling prices collapse back into a swarming state.

In the simulation the data are fed to the agent one by one. It is important to stress that the

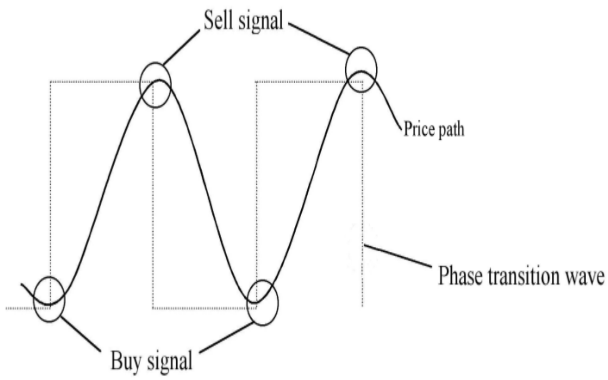

Figure 2: Phase transition wave over a hypothetical price path

phase transition wave is build up as the new data arrive without any recalculation of past positions.

7 Finding the transition point

Technically a measure of the usefulness of a prediction of y from x is the ratio of the sum of squares of the errors of prediction based on the regression line to the sum of squares of those errors when the relationship between x and y is ignored. In that case one would use the average of y as the predicted value of y for all x. This measure is given by

$$\frac{\sum_{i=1}^{n} (y_i - \bar{y})^2 - \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}$$

where $\bar{y}$ is the average and $\hat{y}$ the predicted value. In this form the numerator represents how much the sum of the squares of the prediction errors have been reduced by using regression. It can be shown that the preceding measure is equal to r^2, where r is the correlation coefficient.

$$r = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left[\sum_{i=1}^{n} (x_i - \bar{x})^2 (y_i - \bar{y})^2\right]}}$$

A perfect normally distributed population would be fully covered by the cumulative distribution function. This function can be used to find the relationship between the coefficient and the corresponding fraction of a particular data set. The distribution function is obtained from the Gaussian error function.

$$erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

It has the same relative shape as the cumulative distribution function. The value of the error function at the ordinate x for y between $-\infty$ and x is also an approximation and is evaluated by substituting an infinite series. The new expression is then integrated term by term to produce another infinite series. The result is

$$erf(x) = \frac{2}{\sqrt{\pi}} e^{-y^2} \sum_{n=0}^{\infty} \frac{2^n y^{2n+1}}{1 \cdot 3 \cdot \ldots \cdot (2n+1)}$$

The routine adds a new term to the sum of the previous term. If a particular term does not change the sum by more than a given tolerance, the procedure terminates. The resulting function has a range from 0 to 1. The function is approximately equal to its argument for the range 0 to 0.6. It estimates the probability that an observation will fall within s standard deviations of the mean, assuming a normal distribution.

8 Evaluating the signal

The second part of *Random Harvester* is an expert system. When a trading signal is given the agents looks up his collection of business rules to decide what action to take, i.e. to go long or short, to buy or to sell or to close expiring option contracts. Transactions costs are applied when appropriate. Not every phase transition signal results in a trade.

9 Some Results

The advantage of an trading agent over a human being is his lack of vertigo when riding up or down the trading path. In tests of the agent against a human professional, the human showed signs of hesitation, that in a few cases have led to a better result. *Random Harvester* nevertheless outscored its counterpart in the end. The system improved its results when short positions could be taken. This was only allowed with derived products such as options and futures. The kind of assets used had no influence on the performance: the agent traded equally well with exchange rates, gold, bonds

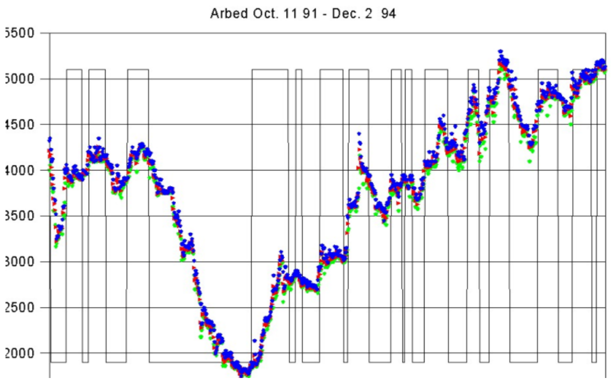

Figure 3: Arbed (Steel Manufacturer)

or stocks. Figure 3 shows the phase transition wave projected over a sample of three years of trading prices for the Arbed stock.

Arbed is an important European steel manufacturer. In the simulation the agent buys and sells the same stock over that period. When sold the asset will eventually be bought back at the right moment. It is a *buy and hold* strategy that combines the normal stock evolution with a *random harvest* of intermediary price differences. This yields with 17 transactions a return of 71% after deduction of the transaction costs. 60% of the return is from *Random Harvester* the remainder is the difference between the initial value and the value of the stock at the end of the experiment.

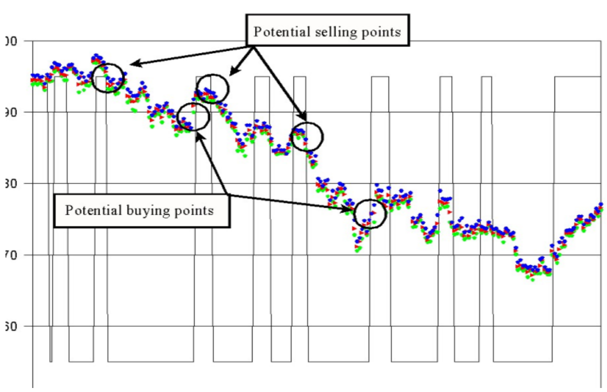

Figure 4: Dollar Guilder Exchange Rate. Jan. - Dec. 94

Figure 4 shows the wave following closely the major up and down swings of the dollar guilder exchange rate over a period of one year at the Amsterdam Stock Exchange. Not every transition point is at the a local maximum/minimum, but the approximation is close enough to allow for interesting trade positions to be set.

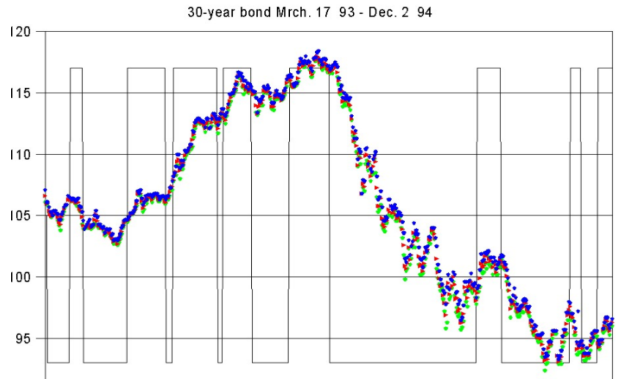

Figure 5: 30-year Dutch bond

In Figure 5 the agent tracks a 30-year bond for almost two years and generates 14 trades with a positive result. With hindsight these were not always the most advantageous moves but no trade was closed at a loss.

Date	**Buy**	**Sell**	**Result**
16 Apr. 93	106,23		
18 Jun. 93	104,76		
30 Jun. 93		106,77	0,54
27 Sep. 93		112,85	8,08
1 Oct. 93	113,04		
3 Nov. 93		114,99	1,95
14 Dec. 93	116,11		
21 Jan. 94		117,51	1,40
27 Jan. 94		116,84	
15 Apr. 94	103,41		
8 Jul. 94	99,11		
4 Aug. 94		101,13	2,01
19 Oct. 94	96,99		
18 Nov. 94	94,26		

10 Conclusion

The trading agent does not try to model every particular price path of an asset. Instead it uses a metamodel that evaluates the data when they arrive. His memory is limited to the current state. Probably there is something to be learned from the lengths between the different phases. It is not clear if an ideal swarm dimension can be estimated for a given asset. The trading agent is not bothered by any fundamental economic insight of the assets he is working with. No comparisons have been made with alternative algorithms. These issues demand further research.

Acknowledgments

My thanks go to the anonymous reviewers for their constructive remarks and also to the *Financieel-Economische Tijd* for providing the data set with asset prices from several European stock markets.

References

W. B. Arthur, S. N. Durlauf and D. A. Lane, eds. 1997. *The Economy as an Evolving Complex System II.* Addison-Wesley. Reading, Mass.

P. Bak, M. Paczuski and M. Shubik. 1996. Price Variations in a Stock Market with Many Agents. Working Paper for the Santa Fé Institute Economics Research Program', submitted to the *Journal of Mathematical Economics.*

P. De Grauwe, H. Dewachter and M. Embrechts. 1993. *Exchange Rate Theory* Blackwell, Oxford.

J. L. Deneubourg, S. Goss, N. R. Franks, A. Sendova-Franks, C. Detrain and L. Chretien. 1990. The Dynamics of Collective Sorting: Robot-like Ants and Ant-like Robots. In J-A Meyer and S. Wilson, eds, *Simulation of Adaptive Behaviour: from Animals to Animats*, MIT Press, Cambridge, Mass.

E.F. Fama and K.R. French. 1992. The Cross-Section of Expected Stock Returns. In *The Journal of Finance*, 2.

K.R. French and R. Roll. 1986. Stock Return Variances, The Arrival of Information and the Reaction of Traders. In *Journal of Financial Economics*, 17.

L. Harris. 1986. A Transaction Data Study of Weekly and Intradaily Patterns in Stock Returns. In *Journal of Financial Economics*, 16.

A. W. Lo and A. C. MacKinlay. 1999. *A Non-Random Walk Down Wall Street.* Princeton University Press, Princeton.

B. Mandelbrot. 1966. Forecast of future prices, unbiased markets and Martingale models. In *The Journal of Business of the University of Chicago*, 39.

B. Mandelbrot. 1998. *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk.* Springer-Verlag, New York.

R.N. Mantegna and H.E. Stanley. 1995. Scaling Behavior in the Dynamics of an Economic Index. In *Nature*, 376.

Hiro-Sato Niwa. 1994. Self-organizing Dynamic Model of Fish Schooling. In *Journal of Theoretical Biology*, 171.

E. P. Peters. 1991. *Chaos and Order in Capital Markets.* J. Wiley and Son, New York.

R. J. Schiller. 1984. Stock Prices and Social Dynamics. In *The Brookings Papers on Economic Activity*, 2.