Theory of Stochastic Process
Vol.0 (00), no.0 ,1999, pp.00-00

THOMAS HELLSTRÖM

ASTA - A TOOL FOR DEVELOPMENT OF STOCK PREDICTION ALGORITHMS

Experiments with a new tool for development of stock prediction algorithms are presented and discussed. Algorithms based on trading rules may be combined with fixed-horizon methods and the evaluation is performed by simulated trading using historical data from the Swedish stock market.

AMS 1991 subject classifications. 90A09.

Key words and phrases. Trading rules, Trading simulations, Returns, Ranks, Stock predictions.

1. INTRODUCTION

This paper reports some experiments with trading rules and time series predictions performed using ASTA, an Artificial Stock Trading Agent written in the Matlab language. The primary purpose of the ASTA system is to provide an easy-to-use environment for research and development of multi-stock trading algorithms. A key component in the system is a stable and realistic test bench, highlighting and helping to avoid the huge risks of data snooping involved in this kind of financial prediction. The behavior of the agent is controlled by buy and sell rules, which can be composed interactively or written as user-defined functions. Various types of combinations of rules and data screening can be easily performed. The system can be used interactively, as a Windows-based development tool, or non-interactively, where a supplied objective function maps a trading strategy to a profit measure. The latter mode can be used to optimize parameters, or automate the development of trading strategies, for example with genetic methods. An example of optimizing parameters can be found in Hellström, Holmström (1999). For a more thorough description of the system, refer to Hellström (1998a). In the present paper the work with development of trading rules is shown in two examples. One example uses a combination of trends and traded volume while the other trading rule is based on a rank measure.

In addition to being a test bench for trading algorithms based on buy and sell rules, pure fixed-horizon predictions can also be developed and tested with ASTA. As an example, an investigation of the correlations between returns for international stock indexes is presented.

2. THE ASTA SYSTEM

In this section we briefly discuss the design and implementation of the ASTA system. The architecture is based on an object-oriented approach with two major objects; the Market and the Trader as shown in Figure 1. The two objects have a number of attributes and operations that can be applied on the objects. The task of the artificial trader is to act on an artificial market with a large number of available stocks that vary in prices over time. The trader has to execute the trading rule at every time step, and decide whether to buy or sell stocks. The sole aim of the trader is to produce as high a profit as possible. Various aspects of the calculation of the performance are discussed in Hellström (1999b). The presentation and evaluation of the trading results are a major part of the ASTA system.

The screen layout for the program is shown in Figure 2. In this picture, the user can set up parameters for a simulated trading. The most interesting items are the lines “Buy rule” and “Sell Rule”. This is where the trading rule is defined. The rules follow Matlab syntax and can include any of the large number of predefined functions or the user's own functions with new algorithms. The simulation of trading starts by clicking on the “Run” button. At each time step the buy and sell rules are executed and acted upon by the artificial trader. The results are finally presented in tabular form as in Figure 2, and in graphs as in Figure 3. The command button “Sweep” initiates a whole series of simulations with different values on symbolic parameters in the Buy and Sell rules. The name of the parameters is given in the “Parameter” text box, and the values to be included in the multiple simulations are given in the “Values” text box. The results are presented in six graphs, also written as encapsulated postscript (eps) files.
Example of this usage can be found in Hellström, Holmström (1999).

3. SIMULATION WITH TRADING RULES

This section presents some examples of the kind of interactive development and testing of trading strategies that can be performed with the ASTA system. 32 major stocks with active trading from the Swedish stock market for the years 1987-1997 have been selected for analysis. In Figure 2, the fields “Buy rule” and “Sell Rule” contain the following rules:

Buy rule:	Trendk(20) < 0 & Gvol10¡0
Sell rule:	Loss>10 | Profit >20

(1)

The gaussian volume function $Gvol10$ is defined as:

$$Gvol10(t) = \ (V(t) - m_V(t))/\sigma_V(t) \tag{2}$$

where $m_V(t)$ and $\sigma_V(t)$ are mean value and standard deviation for the traded volume V computed in a running window $t-9, t-8, ..., t-1$. $Gvol10$ expresses the number of standard deviations by which the volume differs from its running mean.

The k-step trend $Trendk(k,t)$ of a stock m is defined as:

$$Trendk(k,t) = \frac{100}{k} \cdot \frac{Clos(t) - Clos(t-k)}{Clos(t-k)} \tag{3}$$

where $Clos(t)$ is the close price at day t for stock m. By setting k at different numbers we get measures indicating how much the stock has increased per day since its value k days ago.

The system automatically keeps track of the change in price of each stock since it has been bought. The $Profit$ function is defined as:

$$Profit(t) = (Clos(t)\ /\ BuyPrice - 1) * 100. \tag{4}$$

The Loss is sometimes a more convenient measure and is defined as:

$$Loss(t) = (1 - Clos(t)\ /\ BuyPrice) * 100. \tag{5}$$

The trading simulation in the example is set to perform 10 identical runs with a random acceptance of buy signals. A common problem in trading simulation systems is that a huge portion of the generated buy signals are never executed since the trader simply is out of money. By performing multiple simulations this problem is partly overcome. The average result for the artificial trader is presented as annual profits together with the increase in index. The mean difference between these two figures constitutes the net performance for the trader. The performance is displayed in both tabular and graphical formats as shown in the table part of Figure 2 and in Figure 3. The third last line of the table presents the annual profit above index. The mean value of this entity is presented in the second rightmost column and represents a one-figure performance measure. However, the individual figures for each year should also be considered. The results show an average annual profit of 22.4% compared to the 16.3% achieved by the index. We also observe that the trading strategy does worse than the index for 3 out of the total 11 years investigated. A judgement of the significance of the results must also consider the number of trades performed during the simulation. The average number of trades (each buy and each sell counts as a trade) in the example are 62 per year.

We continue with another example where the trading rules are based on a rank concept for the stocks. The k-day returns $R_k(t)$ defined as

$$R_k(t) = 100 \cdot \frac{Clos(t) - Clos(t-k)}{Clos(t-k)} \tag{6}$$

are the primary target in most researches on the predictability of stocks. However, a real trading situation involves not only attempts to predict the individual returns for a set of interesting stocks, but also a comparison and selection among the produced predictions. We therefore introduce a rank concept A_k, based on the k-step return R_k as follows. The k-step rank A_k^m for a stock s_m in the set $\{s_1, ..., s_N\}$ is computed by ranking the N stocks in the order of the k-step returns R_k. The ranking orders are then normalized so the stock with the lowest R_k gets rank -0.5 and the stock with the highest R_k gets rank 0.5. The definition of the k-step rank A_k^m for a stock m belonging to a set of stocks $\{s_1, ..., s_N\}$, can thus be written as

$$A_k^m(t) = \frac{order(R_k^m(t), \{R_k^1(t), ..., R_k^N(t)\}) - 1}{N-1} - 0.5 \tag{7}$$

where the $order$ function returns the ranking order of the first argument in the second argument, which is an ordered list. Thus, $order$ returns an integer between 1 and N. R_k^m stands for the k-step returns computed for stock m. The scaling between -0.5 and $+0.5$ assigns the stock that has the median value on R_k the rank 0. A positive rank A_k^m means that stock m performs better than this median stock, and a negative rank means that it performs worse. This new measure gives an indication of how each individual stock has developed relatively to the other stocks, viewed on a time scale set by the value of k. The scaling around $zero$ is also convenient when defining a prediction task for the rank. It is clear that an ability to identify, at time t, a stock m, for which $A_k^m(t+h) > 0$ means an opportunity to make profit in the same way as identifying a stock, for which $R_k(t+h) > 0$. And even better, since such a method is guaranteed to do better than the average stock. The hit rate for the predictions can be defined as the fraction of times, for which the sign of the predicted rank $A_k^m(t+h)$ is correct. A value greater than 50% means that such a trading strategy would do better than the average stock. The statistical properties of the rank measure have been investigated empirically in more detail in Hellström (1998b). In the following we use the observation that the ranks for a set of stocks exhibit a negative first lag in the auto-correlation function, signifying a mean reverting behavior for the rank measure. We set up the following buy and sell rules:

Buy rule:	Rank5(T)<-0.48	(8)
Sell rule:	Rank5(T)>0.48	

where the ASTA function $Rank5$ for a stock m is computed as $A_5^m(T)$ in definition 7. The buy rule suggests buying if the stock is among the worst performing 2% viewed over the last 5 days. The sell rule suggests selling if the stock are among the best performing 2%. Quite contrary to what is often considered correct, we buy the losers and sell them when they are among the winners. So how does such a strategy perform? In Figure 4 we see the results of 10 trading simulations with random acceptance of the buy rules. The strategy makes 45% annual profit while the index in average increases by 16%. We also observe that the trading strategy is better than the index 8 out of 11 years. The reverse strategy, buying the winners and selling the losers, has also been tested. The results are presented in Figure 5 which shows that the annual profit is only 5% and is better than the index only 2 out of 11 years.

4. FIXED HORIZON PREDICTIONS

ASTA has been primarily developed to evaluate Buy and Sell rules by simulated trading. The other way to evaluate trading rules is a fixed prediction-horizon approach. It is implemented is ASTA in the following fashion. The entity to be predicted is input in the Predict field. The entity that should produce the predictions is input in the Predictor field. The Predictor expression is a function of stock data (close, high , low and volume) available at time T. The Predict expression is typically a function of data available at a time >T. During execution the two expressions are computed, for stocks and times where the Buy rule evaluates to True, i.e. for the occasions when stocks should be bought according to the Buy rule. In this way a selection procedure can be combined with the prediction task. If the Buy rule is left empty, the Predict and Predictor expressions are computed for all steps during the simulation interval. Performance for the predictions is finally output with one line for each stock. The functions entered in the Predictor field can be composed from a large number of predefined functions, which can be easily extended by user-defined Matlab functions that implement prediction algorithms of any complexity.

To illustrate the usage a simple example with correlations between international stock indexes is given. There is a common belief that the Dow Jones index has a great influence on other countries' stock markets. We would like to test and quantify that belief for the Swedish stock market measured by the official stock index Generalindex. The Predict and Predictor fields in ASTA are input as follows:

Predict:	Trendk(1,T)	
Predictor:	Trendk(1,T-1,'DJ')	(9)
Stocks:	SXGEN	

The Predict field is the one-day trend for the investigated security computed at day T. The Predictor field is the one-day trend for the Dow Jones computed at day T-1. The Stocks field is an identifier for the security (Generalindex) for which the Predict field is computed. In Figure 6, the ASTA system for the task is shown for the time period 1989-1997. The results are shown in the lower pane of the screen. The often most interesting performance measures are the total hit rate and the hit rate for the positive and negative predictions separately. The following are the results of the performance calculation:

hit rate	points	+hitrate	points	-hitrate	points
58.74	2002	61.17	1092	55.82	910

(10)

I.e.: an increase in New York is followed in 61.17% of the cases by an increase in Stockholm the following day. A decrease in New York is followed in 55.82% of the cases by a decrease in Stockholm the following day.

The investigation can be extended to inter dependencies between other international stock indexes. The results of such an investigation is summarized in Table 1. The correct interpretation is explained by an example: the sign of the one-day return computed at day T-1 for the American SP500 is the same in 60.7% of the cases as the sign of the one-day return computed at day T for the German DAX index. The largest correlations can be found between the Dow Jones (DJ) / SP500 and the European indexes DAX, SXGEN and FTSE100. The last line shows the performance for a common benchmark: the naive predictor, which simply claims that the index always increases on the next day. Another naive predictor is the constant-return predictor. The performance for this predictor is shown in the diagonal in the table. E.g.: the sign of the return for the Swedish Generalindex is unchanged from one day to another in 57.4% of the cases. The table entries for the Japanese Nikkei50 index are computed on the same day as the indexes to be predicted since the Japanese stock market closes before the other ones open. The predictions are this way be "one day ahead" even in the case with the Nikkei50 index used as predictor.

A selection based on the magnitude of the daily returns can be done by adding a suitable Buy rule to the example shown in Figure 6. E.g.: **Trendk(1,T-1,DJ)>0.5** which generates predictions and performance measures only for days where the daily change in Dow Jones was greater than 0.5%. Table 2 shows the results of this selection procedure. The number of selected points is about 500 for each index (row). As we can see, the correlation between the indexes gets more pronounced as we concentrate on the large positive changes. Example: Days when the SP500 returns are $> 0.5\%$ are followed in 69.6% of the cases by an increase in the Generalindex the following day.

5. CONCLUSIONS

The presented system provides a powerful tool for the development and evaluation of both trading algorithms and fixed horizon stock prediction algorithms. Parameter settings can be tested and data screening can be performed easily interactively. The dangers of "data snooping" get highlighted by breaking down the performance measures into shorter intervals. The inherent uncertainty resulting from the noisy processes involved is partly dealt with by a randomization scheme. Our firm belief is that the major issue when developing trading algorithms and systems is the performance evaluation. Distinguishing a true predicting ability of around 55% from a random coincidence in the testing procedure is a very difficult task that has to be considered seriously. Usage of uniform and integrated test benches such as the presented ASTA system is no doubt the way to go when developing reliable trading systems in the future. The ASTA system is free for academic use. More information can be found on the ASTA home page, Hellström (1999).

BIBLIOGRAPHY

1. Hellström, T., *ASTA - a Test Bench and Development Tool for Trading Algorithms*, UMINF 98.12 ISSN-0348-0542, Department of Computing Science Umeå University, Sweden, (1998a).

2. Hellström, T., *A Random Walk through the Stock Market*, Licentiate Thesis, UMINF 98.16 ISSN-0348-0542, Department of Computing Science Umeå University, Sweden, (1998b).

3. Hellström, T. and Holmström, K., *Parameter Tuning in Trading Algorithms Using ASTA*, In Abu-Mostafa, Y. S., LeBaron, B., Lo, A. W., and Weigend, A. S., editors, *COMPUTATIONAL FINANCE*, 1999, MA: MIT 3. Press, New York, Cambridge, (1999).

4. Hellström, T., *ASTA Home Page*, `http://www.cs.umu.se/~thomash/asta/home`, (1999).

Department of Computing Science, Umeå University, S-90187 Umeå, Sweden
E-mail: thomash@cs.umu.se
http://www.cs.umu.se/~thomash

	DJ	SP500	Nikkei	DAX	SXGEN	FTSE
DJ	50.5	51.0	53.9	58.8	58.7	55.0
SP500	53.7	51.3	55.3	60.7	58.9	56.7
Nikkei	54.0	53.8	50.1	50.4	51.8	49.0
DAX	49.6	49.9	51.7	50.6	52.4	49.6
SXGEN	50.5	51.4	53.1	53.9	57.4	49.8
FTSE	50.0	51.5	53.9	54.1	54.1	50.1
Naive	55.0	54.8	49.3	53.5	52.5	52.9

Table 1: Correlation between international indexes for 1987-1997. Each figure is the fraction with equal sign of the one-day returns for pairs of indexes. The indexes on the rows are computed at time T and the columns at time T+1.

	DJ	SP500	Nikkei	DAX	SXGEN	FTSE
DJ	56.7	55.5	58.4	69.6	70.8	65.9
SP500	61.4	56.7	57.1	73.8	69.9	68.7
Nikkei	60.1	59.4	48.4	53.1	52.6	50.2
DAX	51.6	52.2	52.4	53.8	53.5	51.1
SXGEN	56.3	55.7	53.7	56.4	62.8	53.5
FTSE	55.2	55.8	56.4	61.2	59.8	53.5
Naive	55.0	54.8	49.3	53.5	52.5	52.9

Table 2: Same as Table 1 but with a selection of points where the daily returns for the predictor (the row labels) are >0.5%.

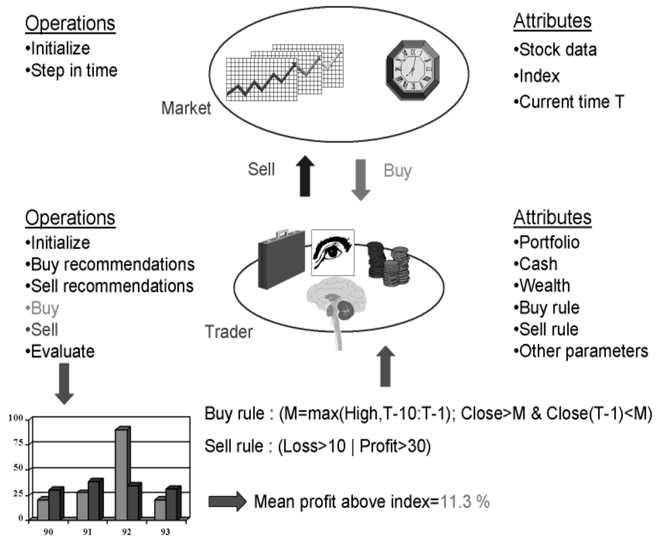

Figure 1: Architecture of the ASTA system. The artifical trader executes the Buy and Sell rules at every time step and trades stocks on the market with historical data.

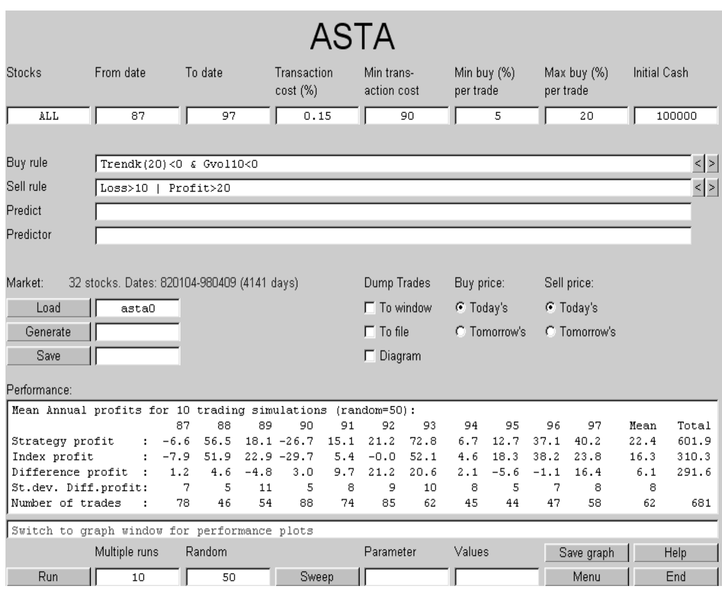

Figure 2: ASTA setup for 10 simulations with Buy rule=Trendk(20)<0 AND Gvol10<0 and Sellrule=Loss>10 OR Profit>20

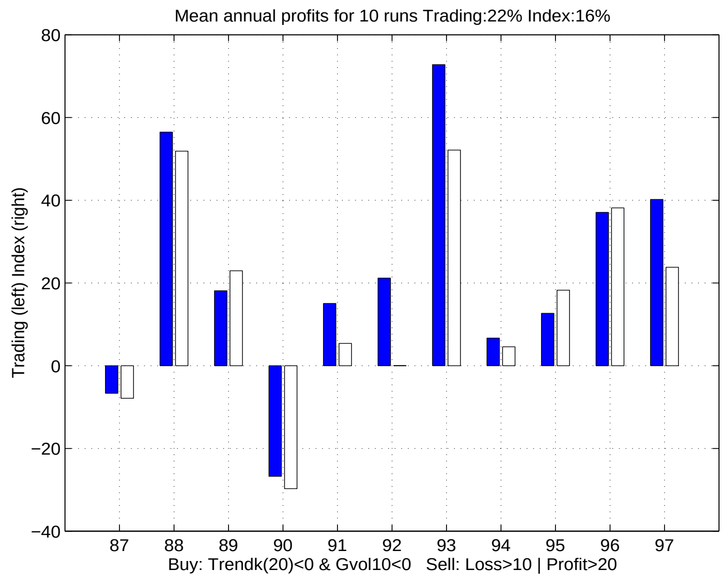

Figure 3: Graphical results from the simulations in Figure 2.

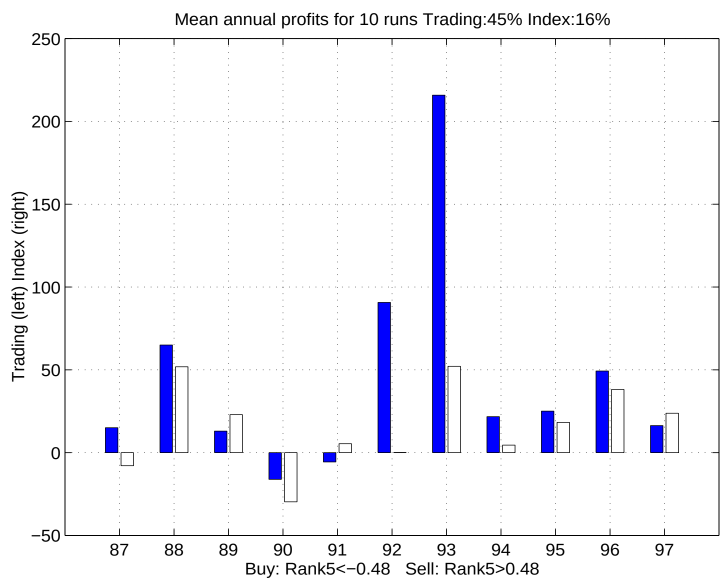

Figure 4: Mean trading results from 10 simulations. Buy the losers, sell the winners. The annual result is better than the index for 8 out of 11 years.

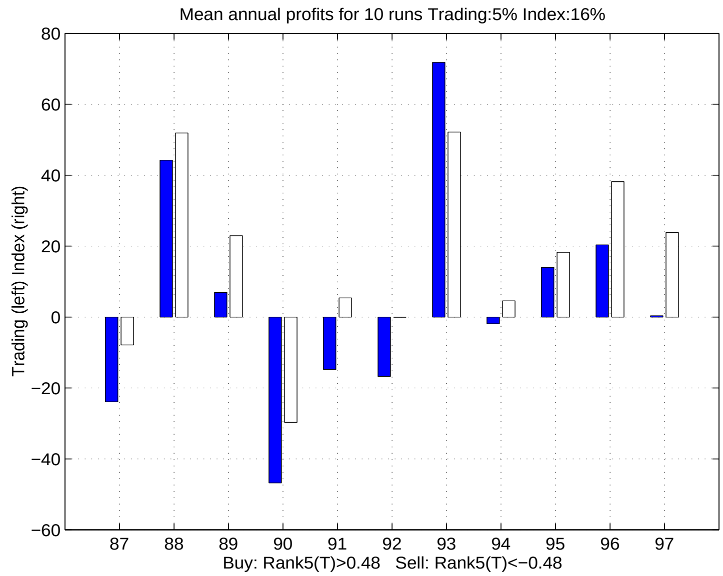

Figure 5: Mean trading results from 10 simulations. Buy the winners, sell the losers. The annual result is better than the index for 2 out of 11 years.

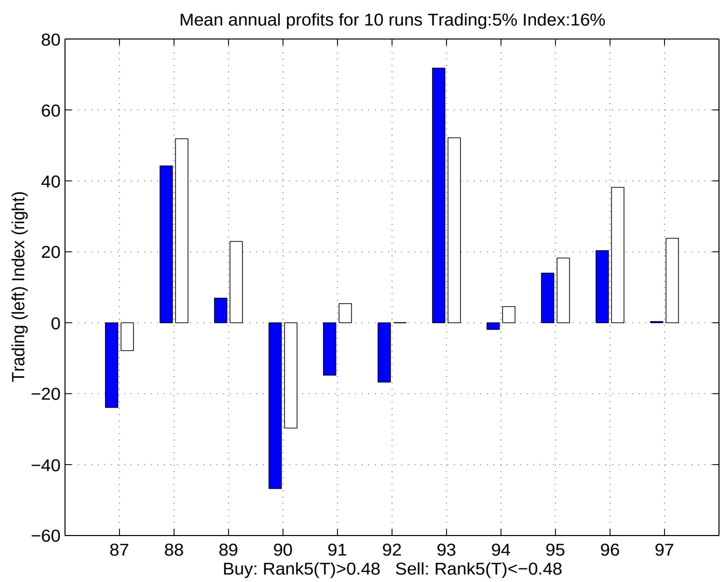

Figure 6: Time series predictions with ASTA. The one-day return for the Swedish Generalindex is predicted as yesterday's one-day return for the Dow Jones index. The hit rate is 58.74%.