16

Genetic Algorithm Application for Trading in Market toward Stable Profitable Method

Tomio Kurokawa
Aichi Institute of Technology,
Japan

1. Introduction

One application of a training system, including genetic algorithm (GA), to trading in technical market is to create a good model trader using past market data so that the model trader makes good performance in unseen future market data space. It is usually easy to create a model trader to work well for in-sample data. However, it frequently happens that the created model trader makes poor performance for unseen data. This phenomenon is termed as overfitting. This chapter provides some insight into overfitting in the environment of trading in market — enormously wide spaces of in-sample, out-of-sample data and technical model trader's space — and proposes some solution coping with the problems which exist in the spaces.

2. Genetic algorithm as learning system

Learning has been considered as one of the most powerful problem solving techniques for technical market analysis or trading in market. Among many learning systems so far we have devised, genetic algorithm (GA) is considered as one of the most powerful.

Using GA as a learning system for technical market, there are three spaces (Geurts, 2005) we have to take into consideration. The first one is the past data space from which we obtain in-sample data and with which we train the market evaluation systems, say, model traders. The second one is the space of model traders. It could be a simple one—small model space to a complex model trader—large model space. The third one is the unseen future data space to which we apply the trained model trader and hopefully obtain preferable results.

All of the three spaces are so wide compared with the data we can use as past data or the data we will see in the future. As for the past data space, data for a company we can use is only an instance of the data space. We cannot have more than one instance, just one. For unseen future data, it is also true that the one we will see is just one instance, no more than that. The model trader space is possibly called as technical indicator space. The number of algorithm we can try is very limited. Computer generated ones, which are not limited, are inclined to become too particular, lacking of comprehensibility and resulting in overfitting.

3. Background literature

Overfitting has been the major problem of learning of technical market decision. Numerous researches have been done using evolutionally methods (including GA) in the literature. In some past studies, solutions were tried to be sought. Those are categorized into three areas as noted below.

First is about in-sample data selection. Wang et al. employed, in GA, the methods with which newer in-sample data were used for training by sliding the in-sample data window as the trading proceeds (Wang & Chen, 1998) . Lam et al. employed sliding (incremental and dynamic) in-sample data approach for training the system with GA and fuzzy mechanism (Lam et al., 2002). These schemes were based on the presumption that newer data might be better representing coming unseen out-of-sample data. Neely et al. used, in GP (genetic programming), a technique regarded as validation procedure, in which a selection period is placed after the training period in order to select one good program for next generation (Neely et al., 1997). Kurokawa demonstrated that trading chance was increased and at the same time overfitting was reduced by increasing number of stock names concurrently monitored (Kurokawa, 2008, 2009, 2011).

Second is related with learning itself. It is to improve the learning process so that trading rule should have good performance and avoid overfitting. Obtaining a good performance rule is concerned with many things such as learning process itself, elements of computational structure including indicators, fitness strategies, maintaining simplicity, generality and many others. In the past, Becker et al. using, in GP, reduced set of operators, set of increased indicators (elements of computational structure), and complexity penalizing strategy (fitness strategy) in the training process maintained the simplicity of the tree structure of GP-generated rule by limiting the number of nodes and the depth of the tree (simplicity) (Becker & Seshadri, 2003a, 2003b) . Simplicity plays very important roles for avoidance of overfitting because complexity is more likely to cause overfitting (Becker & Seshadri, 2003a). Lin et al. set sub-ranges for parameters of technical trading rules (fitness strategy) by GA and obtained robust results (Lin et al., 2005).

Third is to devise effective technical indicators. Technical indicators play important roles in rule making with learning, especially in evolutionary process. In the process, they are usually pre-given as essential components. For GA, indicators with computational structure are given and the parameters are optimized. For GP, though it has the ability to find good computational structure of technical rules, indicator functions such as moving averages, etc. must be given as basic component at the beginning. Eventually, some of the components play important roles in the generated rules. Such indicators were also used in GP process of the studies by Becker et al. (Becker & Seshadri, 2003a) and they showed positive results. Potvin et al. applied GP to Canadian individual stocks and reported that it did not necessarily outperform buy and & hold (B&H) approach (Potvin et al., 2004). Pavlidis et al. compared moving average based rules and GP-generated rules on money exchange rates and obtained the results that both are profitable but moving average based rule is more robust than the GP-generated one (Pavlidis, et al. 2007) . Mabu et al. included several conventional indicators in GNP (genetic network programming) and showed promising results (Mabu et al., 2007). Kurokawa tried to seek better technical indicators used in GA process (Kurokawa, 2007). More recently, Lohpetch et al. showed a method which gains fairly robust generation of trading rules which outperforms B & H (Lohpetch et al. , 2010) using GP and monthly trading.

4. Potential areas for study

Looking over a stock market, the data are very vast. Even data of a single stock name are vast when considering unseen out-of sample data of potentially numerous patterns. This makes it unreasonable to try to cope with it by learning only a small portion of past data. One or two year market data of a stock name are very small compared to unseen potentially vast out-of-sample data. Learning process is to make the target system particular. It is to adjust the system to the in-sample data. Accordingly, the target system hardly becomes general.

In order to solve the problem, some generality introducing mechanism is necessary. Possible solutions may exist in 1) in-sample data of large size and concurrent monitoring of large set of stocks in trading, 2) sophisticated but simple learning process which can make use of large set of stocks and avoidance of adjusting to the particularity of in-sample data or overfitting and 3) devising effective technical indicators.

For the first potential solution, most studies in literature about market timing by learning have been done with in-sample data of small size, that is, one or two years for a single name or an index or of several years. Small size data usually cannot have generality. Naturally it is hardly possible for any process to extract generality from data of small size. Hence in-sample data of large size are needed. Larger size data are more likely to have more generality. However, they are also difficult to handle and time consuming for processing. There are two ways of expanding in-sample data. One is with data of long period. There are, however, some limitations about the size. The other is a larger set of stock names to concurrently monitor. The latter case was examined with GA procedure using data of hundreds of stock names (Kurokawa, 2008, 2009, 2011). In this study, it is more extensively examined.

For the second potential solution, this is the area where many studies have been done in literature. In order to handle data of large size, effective learning mechanism with fast processing speed and ability avoiding overfitting becomes important. Both of sophistication with simplicity and speed are necessary at the same time. The learning should have the ability to handle the data of large size and to extract the generality. In this study, this concept was examined by introducing indirect fitness control with profit related indexes, which is supposedly essential for stable profit.

The third potential solution is concerned with technical indicator, which is directly related with data computation. It is the device by which trading signals are computed and detected. Since many market timing systems with learning employ technical indicators, developing effective indicators is essentially important. What's more, indicators are independent of learning process in the sense that they are made before learning process operates. Hence, they could bring about generality if not totally influenced by learning. Simplicity and comprehensibility could be given by human heuristics. Hence, effective indicators are of great utility. Some studies were made in this area (Kurokawa, 2007). In this study, however, a technical method (Takizawa, 1999) was used.

5. Three step experiments

This study is of three steps for trading performance improvements and its experimental demonstrations. The first one is to investigate what is overfitting and how it is related with trading performance. The second one is to demonstrate the first one by stock switching. The third one is to show some ways of seeking generality of stable and profitable trading with extensive experiments.

6. First step experiments

To examine how influential the selection of in-sample data is, three experiments using GA with the market timing indicator SP-method (Takizawa, 1999; Kurokawa, 2009) were conducted with different stock data of different sizes but of the same data period. Data from Jan. 1, 2001, to Dec. 31, 2002, were used for training and those from Jan. 1, 2003 to Dec. 31, 2004 for testing. These experiments were mainly to get information on market features, not just to find the best trading model. It did not calculate trading returns for a period but certain profit related indexes. In the experimental trading, a unit of stock was bought whenever a buy signal appeared and when the profit rate exceeded a predetermined level or stock holding went beyond a predefined number of days, the stock was sold. Trading was independent of the amount of cash held.

Another important point here is that the trading did not depend on the timing of incidentally selected buying/selling, so these results show the features of the entire stock market better. Ordinary trading simulation tends to depend on timing. Some buy decisions for some stock mean not buying other stocks — results by these decisions are dependent on timing.

Chromosome was composed of a set of parameters specifying how to run trading and was optimized in GA. They were Gene0, Gene1, Gene2, Gene3, Gene4, and Gene5 as:

1. Gene0: SP-wave rate, SP%
2. Gene1: SP-minus change rate, SP%-
3. Gene2: maximum number of days to hold bought stock
4. Gene3: minimum recovery rate to sell bought stock
5. Gene4: minimum falling speed of price
6. Gene5: maximum price level

6.1 How GA experiments were done

An ordinary fitness criterion in stock trading is conceptually simple — maximizing profit. The experiments here, however, calculate more than mere profit; they calculate the following profit-related indexes:

- Total profit (TP): the sum of all unit trade profit. "Trade" is used interchangeably here as a stock unit buy/sell pair.
- Winning count (WCT): the count of profitable trade
- Win rate (WRT): the profitable trade count divided by all trade count
- Total of profit rates (TPR): the sum of profit rates on all trades
- Average profit rate (APR): the total of profit rates divided by the number of trades.

TPR, a special index, usually not used, was to get overall market features of both quality and quantity of trading, which is why it was selected for fitness. Some quality (profit rate) must be maintained because transaction cost was ignored in the experiments.

The following represent specific definitions:

$$profit(i) = sell\_price(i) - buy\_price(i), \tag{1}$$

$$TP = \sum_{i=1}^{N} profit(i), \tag{2}$$

$$PR(i) = \frac{profit(i)}{buy\_price(i)}, \tag{3}$$

$$TPR = \sum_{i=1}^{N} PR(i)\ , \tag{4}$$

where i is the identifier for each trade, *buy_price(i)* the stock price at the buying of trade i, *sell_price(i)* the stock price at the selling of trade i, and N total trade count. *TPR* was used as fitness to give each trade equal weight to make it independent of individual stock name's price levels. *TP* might be greatly influenced by some stocks with very high prices.

An ordinary GA was used for optimization, as shown in Fig. 1. GA parameters were crossover rate: 0.7, mutation rate: 0.1, population size: 20, number of generations: 200 and elitisms. The population size was set rather small to speed up large-scale data processing but within a generally allowable range. These specific parameters were chosen arbitrarily and involve no particular reasoning. The parameters of population size 20 and number of generations 200 were rather small, so hitting a globally optimal point is not necessarily expected. Limited optimization, however, is considered acceptable here. Experiments 1-3 are detailed below.

6.2 Data size and overfitting

6.2.1 Experiment 1-1

Ten trading simulations were done independently, one for each of ten stock names, with training done using in-sample data for each stock name, followed by testing using out-of-sample data for the same stock name. Results are shown in Table 1. WCT, TP, and TPR values in testing were much lower than those in training. Stock names No. 2, 4, 6, 7, and 8 gave no chance for trading in testing. "N/A" in Table 1 suggests that the system looked for trading signals but in vain — typical for overfitting. The model trader was so adjusted to in-sample data that it could not find a chance for out-of-sample data. The average successful trade count was 4.7 in training but only 0.7 in testing. The one single chance for Asahi Glass, for example, was not successful in testing.

Experimental results suggest that the model trader trained with a small segment of stock data could lose chances for trading. The target obtained with small in-sample data may thus not be suited to unseen data, which is assumed to be just one incident out of a potentially vast number of data patterns.

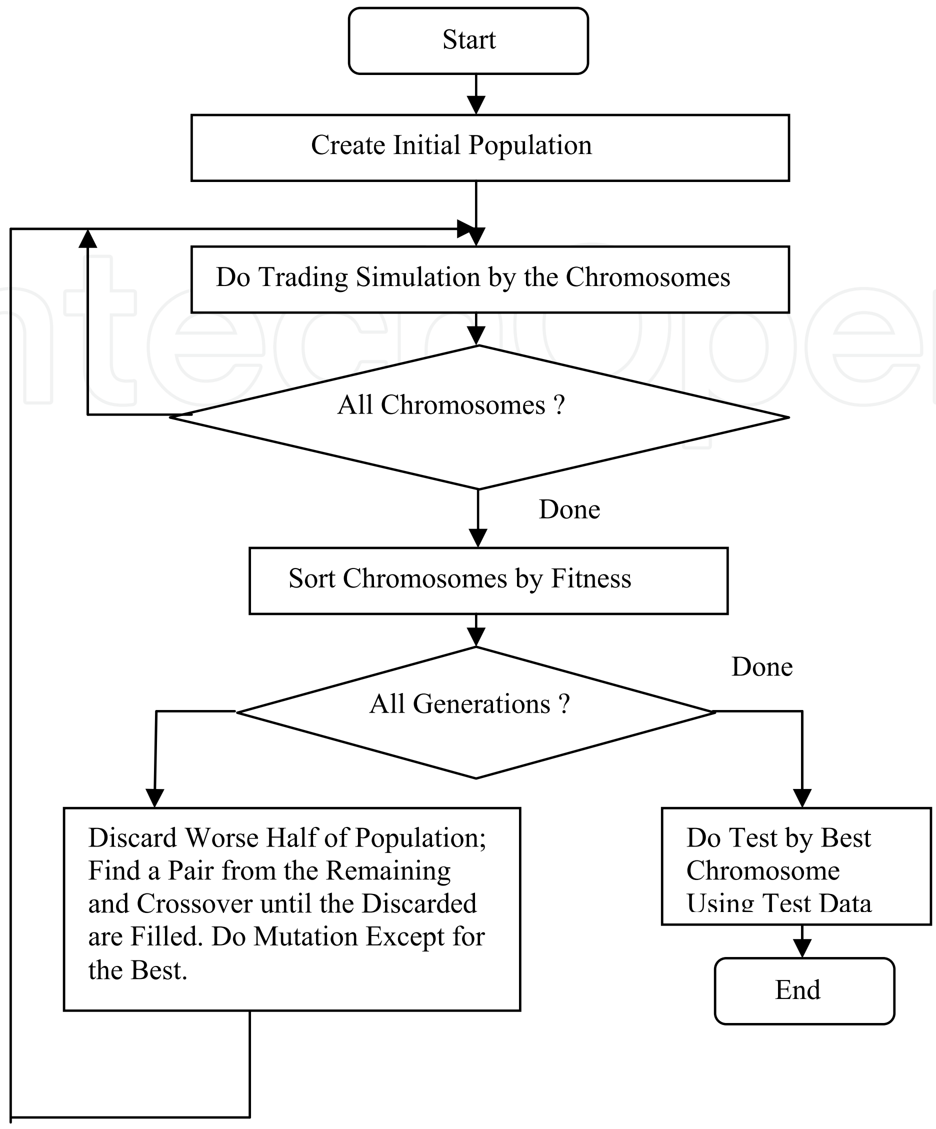


Fig. 1. GA Procedure

	Training (2001–2002)					Test (2003–2004)				
Stock Name	WCT	TP	APR	WRT	TPR	WCT	TP	APR	WRT	TPR
1 Shimizu	4	386	0.223	0.800	1.117	1	151	0.429	1.000	0.429
2 Itoham Foods	11	306	0.073	0.846	0.947	0	0	0.000	N/A	0.000
3 Oji Paper	3	456	0.316	1.000	0.948	1	84	0.180	1.000	0.180
4 Sumitomo Chemical	2	260	0.372	1.000	0.745	0	0	0.000	N/A	0.000
5 Asahi Glass	5	865	0.212	0.833	1.271	0	−20	−0.030	0.000	−0.030
6 Sumitomo Heavy Ind.	2	44	0.355	1.000	0.710	0	0	0.000	N/A	0.000
7 Toshiba	4	401	0.244	1.000	0.975	0	0	0.000	N/A	0.000
8 Matushita Electric Works	4	173	0.062	1.000	0.249	0	0	0.000	N/A	0.000
9 Sumitomo	6	788	0.147	0.750	1.174	3	159	0.100	1.000	0.299
10 Yusen	6	347	0.152	1.000	0.915	2	109	0.127	1.000	0.254
average	4.7	402.6	0.216	0.923	0.905	0.7	48.3	0.081	N/A	0.113

Table 1. Results by training using past data for one stock name for each line (10 names)

6.2.2 Experiment 1-2

Experiment 1-2 used the same stock data of those names as for Experiment 1-1. By concurrently monitoring stocks for 10 names for the in-sample period, just one trading rule was generated instead of one rule for each name. The specified fitness was TPR, the same as that for Experiment 1-1. Results are shown in Table 2. The same procedure was executed 10 times, and the averages were computed and are shown at the end of the table.

In experiments, trading opportunities appear to have been reduced in testing, apparently the result of overfitting, or losing opportunities. Overfitting, however, appears to have been reduced in these experiments. The quality of trading (APR and WRT) decreased both in training and testing from Experiment 1-1 to Experiment 1-2. Good trading opportunities in Experiment 1-2 (WCT: 37.3 in training and WCT: 13.5 in testing), however, were much improved over those in Experiment 1-1 (WCT: 4.7 in training and WCT: 0.7 in testing) and overfitting appears to have been somewhat reduced in Experiment 1-2 from the viewpoint of trading opportunities. Both APR and WRT were slightly lower in testing than in training in Experiment 1-2. APR of 0.064 and WRT of 0.743 in testing appear very good compared to the results (He et al., 2007) — return of 0.0212 for GP1, return of -0.1792 for GP2, and returns by others for Japanese stocks of the same period, though the comparisons are not direct. It may not be decisive, however, for the Japanese stock index Nikkei average, which increased about 30% during 2003 to 2004. Stocks of the ten names examined in the experiments actually had 78.5% gain as average during the period.

APR and WRT in Experiment 1-1 and 1-2 are difficult to evaluate, but given the results of "N/A" (No Trade), trading as done in Experiment 1-1 cannot be said to be advantageous.

Note that trading signals or patterns should have appeared in many places among stocks for multiple stock names.

	Training (2001–2002)					Test (2003–2004)				
Exp. No	WCT	TP	APR	WRT	TPR	WCT	TP	APR	WRT	TPR
1	23	1160	0.117	0.885	3.052	8	270	0.081	0.800	0.813
2	23	1160	0.117	0.885	3.052	8	270	0.081	0.800	0.813
3	71	1209	0.024	0.607	2.817	22	546	0.052	0.759	1.512
4	44	1298	0.060	0.830	3.186	11	173	0.038	0.733	0.576
5	46	1346	0.052	0.793	3.036	15	330	0.042	0.750	0.835
6	48	1278	0.049	0.814	2.877	16	270	0.045	0.762	0.950
7	56	1587	0.033	0.589	3.141	32	728	0.042	0.653	2.069
8	16	1304	0.112	0.667	2.681	4	208	0.075	0.444	0.673
9	23	1158	0.120	0.920	3.003	9	331	0.099	0.900	0.987
10	23	1160	0.117	0.885	3.052	10	355	0.081	0.833	0.976
average	37.3	1266.0	0.080	0.788	2.990	13.5	348.1	0.064	0.743	1.020

Table 2. Results by training simultaneously using past data for 10 stock names

6.2.3 Experiment 1-3

Experiment 1-3 used stock data of 844 stock names, from which those priced above 2,000 Yen were discarded, from Tokyo Market Division I. Experiments were done under the same condition of Experiment 1-2 except the number of stock names and the price requirement. Stocks for all names were monitored concurrently. One single trading rule for stocks of all names was generated using in-sample data instead of generating one rule for each stock name. Results are shown in Table 3. Specified fitness, TPR, was the same as that in Experiment 1-1 and 1-2.

WCT, TP, and TPR values increased greatly in both training and testing, suggesting that opportunities for trading increased in Experiment 1-3. Although not shown explicitly, it appears that there should have existed many stock names for which no trading signals appeared in the testing period. Average APR in training declined from Experiment 1-2 to Experiment 1-3 from 0.080 to 0.037, although the reason remains unclear. It may conceivably have been caused by price declines for a large number of stock names. Average APR values increased in testing from 0.064 in Experiment 1-2 to 0.071 in Experiment 1-3 and from 0.037 in training to 0.071 in testing in Experiment 1-3, possibly due to many successful signal detections in many places in the out-of-sample period.

The overall results of Experiment 1-3 seem to suggest that overfitting was reduced greatly by the concurrent monitoring of numerous stocks. Attention should therefore be paid to the results showing that differences in WCT, TP, and TPR between training and testing decreased as data size increased. As stated earlier, Experiment 1-3 also shows that signals caught by the model trader appeared in many places among stocks of multiple names, as is shown by the large WCT, TP, and TPR values in Table 3.

APR average of 7.1% and WRT average of 70.6% in testing appear to be very good compared to the study results (He et al., 2007) as shown in Experiment 1-2, though the comparisons are not direct. It also appears convincing because the number of stock names used in the experiment is large. It is not actually decisive, however, because the Nikkei Average during 2003 to 2004 rose about 30%, very large gain.

	training (2001-2002)					test (2003-2004)				
Exp. No.	WCT	TP	APR	WRT	TPR	WCT	TP	APR	WRT	TPR
1	1309	35810	0.051	0.607	109.627	799	14764	0.060	0.662	72.342
2	2696	42040	0.023	0.568	109.298	1769	35218	0.041	0.631	115.556
3	2989	45648	0.025	0.628	121.000	1948	33923	0.042	0.706	117.332
4	359	14381	0.112	0.634	63.217	21	828	0.307	0.875	7.360
5	4498	48310	0.018	0.639	126.864	2940	46511	0.034	0.726	139.323
6	1330	37274	0.050	0.608	109.137	847	16186	0.059	0.664	75.383
7	4183	45704	0.018	0.611	123.665	2742	45195	0.036	0.699	141.175
8	1859	38295	0.035	0.601	107.644	1264	26552	0.055	0.685	101.528
9	4774	45915	0.016	0.630	123.723	3170	50035	0.033	0.720	146.309
10	2912	47148	0.025	0.606	118.826	1906	34726	0.044	0.688	120.554
average	2690.9	40052.5	0.037	0.613	111.300	1740.6	30393.8	0.071	0.706	103.686

Table 3. Results by training simultaneously using past data for 844 stock names

6.3 Quantitative comparison of the results and overfitting phenomenon

Overfitting is the phenomenon in which poor performance appears for out-of-sample data despite good performance for in-sample data. We define this somewhat more formally as the ratio:

$$Overfitting = \frac{PerformanceForInSampleData}{PerformanceForOutOfSampleData} \tag{5}$$

Table 4 shows the experimental ratios. The figures in Table 4 were calculated using averages given at the bottoms of Tables 1 to 3. Note the changes in ratios during the transition in data size. TPR was used as fitness, so primary attention should be paid to the related aspects. Values in TPR row no. 3 are these ratios, which decrease as data size increases. For small-scale data, TPR performance for the in-sample was 8.01 times greater than that for the out-of-sample. For medium-scale data, it was 2.93, becoming 1.07 for large-scale data. This means that the performance difference between in-sample and out-of-sample data decreases as data size increases. This can be called reducing of overfitting by larger-scale data. Performances other than TPR showed similar trends. Take WCT, for example. Winning signals caught in the in-sample period appeared in the out-of-sample period more often for large-scale data but not for small-scale data.

Smaller ratios are seen for larger-scale data for APR and WRT also. These, however, should be regarded as accidental. The ratios smaller than 1.0 indicate that performance for out of sample data could be better than that for in-sample data, depending on the situation.

We believe that the performance difference between in-sample and out-of-sample periods should not be too big. Any big difference is usually considered caused by overfitting. If the difference is too big, the system is generally considered as not working well, possibly unstable.

Ratio: In-Sample / Out-of-Sample	One Stock Name: Small-scale Data (Experiment 1-1)	10 Stock Names: Medium-scale Data (Experiment 1-2)	844 Stock Names: Large-scale Data (Experiment 1-3)
1. WCT	4.7/0.7=6.71	37.3/13.5=2.76	2691/1741=1.55
2. TP	402.6/48.3=8.34	1266/348=3.64	40053/30394=1.32
3. TPR	0.905/0.113=8.01	2.990/1.020= 2.93	111.300/103.686=1.07
4. APR	0.216/0.081=2.67	0.080/0.064=1.25	0.037/0.071=0.52
5. WRT	N/A	0.788/0.743=1.06	0.613/0.706=0.87

Table 4. Performance comparison between in-sample and out-of-sample data and between data sizes

6.4 Other aspects of overfitting

As stated before, overfitting is the phenomenon in which poor performance appears for out-of-sample data despite good performance for in-sample data. It is usually not difficult to obtain a good model trader for in-sample data. It means that the model trader space is usually wide enough so that training system can find a good performance model trader for the sample data. Many of training systems such as genetic algorithm, neural network (NN), or genetic programming (GP) usually works well for creating a good model trader for in-sample data. NN with two or more hidden layers or GP with much flexibility of programming usually have such a wide space for model traders that it is inclined to produce a kind of nonsense trader which often makes poor performance for unseen out-of-sample data. GA is not an exception. It often produces poor performance traders for out-of-sample data. It often depends on what to optimize. In addition, the created trader with NN or GP usually does not have comprehensibility in the trading rules of the trader.

However, GA which is usually used to adjust a model trader, already structured, to the in-sample data. Owing much to the past researchers of technical analysis, GA is used to find good parameters for the model traders. The technical model traders have spaces wide enough that GA can easily adjust them for the model to work well for in-sample data. However, it is not usually easy for the created traders to work well for unseen out-of-sample data. Overfitting also often appears for GA trained model traders. This overfitting is one of the biggest problems many researchers have tried to solve for market trading. However, it is expected to be smaller for GA trained technical traders than for those by GP or NN. It is probably because the model traders' space for GA is much smaller and because their computation structure is often well organized. This is supported by the study (Pavlidis, et al. 2007).

7. Second step experiments: Switching from one stock to another with data of different sizes

It has been pointed out that large size data have more generality and therefore possibly more effectiveness than a small size data to learn. In other words, there are more generality in training and more opportunities in practical trading. However, large size data are not easy to handle. Based on the demonstrated effectiveness by the experiments with large size data (Kurokawa, 2009), simulations in this study were organized to examine the model trader of switching from one stock to another. In order to see the essence, the model was designed as simple as possible.

Two kinds of experiments were done with different stock data of different sizes for comparison. The experimental procedure was almost the same as the procedure shown by Fig. 1, which was used in the first step experiments except trading was with stock switching. They were the following two:

1. Experiment 2-1 with each single stock name (actually no stock switching), and
2. Experiment 2-2 with 10 stock names for concurrent monitoring.

But, the data periods were the same between the above two and also the same as the first step experiments. The data from 2001/1/1 to 2002/12/31 were used for training and those from 2003/1/1 to 2004/12/31 were for testing. All data were from daily data of Tokyo Stock Market Division I.

The following are how the experimental trading was done. At the start, some amount of cash, supposedly very large, was provided. For the experiment with a single stock name, whenever buy signal appeared, the stocks of the name were bought as many as possible with the available cash and when the sell condition appeared (when the profit rate became more than predetermined level or stock holding length got beyond the predefined number of days), all the stocks were sold. It is a very simple model trader with buy first and sell.

The model trader employed the same chromosome in structure and function which was used in the first step experiments. They are the genes of Gene0, Gene1, Gene2 and Gene3, Gene4, and Gene5 as previously stated. For the model with concurrently monitoring multiple stock names, more than one name could show buy timing at same time. In the case, the stock name with the smallest SP-minus change rate was chosen.

7.1 Computed indexes, fitness, and GA parameters

The first step experiments were for surveying the market — how trading signals were distributed among stocks of multiple names, how quantity and quality of trading were, etc. The second step experiments were to do more practical trading simulations taking advantage of the results of the first step experiments. The organized model trader was to do trading with stock switching among multiple stock names taking advantage to increased trading signals distributed among them and decreased overfitting. The purposes of the second step experiments are to examine how profitable the stock switching is and to compare the results with those of without stock switching.

In stock trading, an ordinary criterion for fitness is simple. It is to maximize the profit or the return. Since the purposes of this step experiments are different from the first ones, the indexes to be computed were changed from the first. The total return (*TRN*), the amount of cash at the end of the simulation divided by the initial cash at the start of the simulation, was used as fitness. Some other indexes were also computed to see how the experiments proceeded. They are the number of trades (*N*: total count of trades), win count (*WCT*: count of profitable trades), win rate (*WRT*: *WCT* divided by *N*), average return (*ARN*: total of individual return divided by the number of trades). "Trade" is used as a pair of buy and sell. The following equations are given for the specific definitions:

$$RTN(i) = \frac{sell\_price(i)}{buy\_price(i)}, \tag{6}$$

$$ARN = \frac{1}{N}\sum_{i=1}^{N} RTN(i)\,, \tag{7}$$

$$WRT = \frac{WCT}{N}\,, \tag{8}$$

where *i* is the identifier for each trade, *buy_price(i)* is the price at the buy of trade *i*, *sell_price(i)* is the price at the sell of trade *i* and *N* is the total count of trades, *WCT* is the number of trades with plus profit and *TRN* is the total return, that is, the final amount of cash divided by the initial amount of cash. The actual computation was done as Equations (6) through (9).

$$TRN = \prod_{i=1}^{N} RTN(i) \tag{9}$$

As seen by Equations (6) through (9), transaction costs were ignored.

As for the optimization process, the same GA procedure and the same GA parameters were used as in the first step. They were crossover rate: 0.7; mutation rate: 0.1; population size 20; generation length 200 and elitisms. The same discussions concerning the optimization may be possible as in the first.

The population size was rather set to the small number for fast processing of large size data but within a generally allowable range. Those specific parameters were arbitrarily chosen and there is not a particular reason for them. Since the parameters of population size 20 (rather small) and generation length 200 (rather short) were also arbitrarily chosen, hitting global optimal point might not necessarily possible. The very best optimization was not supposed for this study.

7.2 Switching strategy

In the trade model of switching from one stock to another, the selection policy for next stock to buy becomes very important when more than one stock shows trading signals. There might possibly be many ways to select next stock. What is most effective is yet to be studied and is left for future study. However, a very simple but probably good method was picked up but somewhat arbitrary in this study. That is, the stock with the smallest SP-minus wave rate was selected from among the stocks showing buy timing. The very small values (minus) of SP-minus rate were supposed to show a bottom of stock price like moving average method but somewhat differently.

7.3 Experiments

The details of the two experiments were as follows.

7.3.1 Experiment 2-1

In the experiments, no switching strategy could be employed, because only one stock name was used for each simulation. Ten trading simulations were done independently, one for each of ten stock names. They are the ten used in the first step. In each simulation, training was first done by using in-sample data of each stock name, and then test was made using the out-of-sample data of the same name. The results are shown in Table 5. N-No. column shows stock name number as given previously in Table 1. As seen, the values of WCT and N in testing are much lower than those in training. On the lines with N-No. 6 and 8, there are symbols of "NA" which mean there were no opportunities for trading in the test period. These results seem showing typical overfitting. The average counts of successful trade were 5.7 in training and those in testing were only 1.7. However, the most important result, the return, TRN for the out-of-sample period was 1.207 as the average. It is not a bad result. Additionally, the win rates (WRT) also were surprisingly good, i.e., 0.921 for in-sample period and 0.811 for out-of-sample period.

At the right end of the table, start and end prices of each stock are shown as well as each return by buy and hold strategy (BHRTN). Start-P is price at the start of the out-of-sample the period and End-P is the price at the end of the period. On the bottom of BHRTN is the average of BHRTNs. It is 1.785. The average return of Experiment 2-1 is 1.207. This value

1.207 seemed good until the average return of the ten stocks turned out to be 1.785. Nikkei average during 2003 to 2004 increased about 30%. Using the start and end prices of No. 6 stock, the return of buy and hold becomes 5.69, extremely high value. The price of the stock rose up so steep in the period. Without the stock, the average BHRTN of the remaining nine is about 1.351. It is still far better than the average return of this experiment 1.207.

Experiment 2-1													
	In-Sample Period					Out-of-Sample Period							
N-No.	WCT	N	WRT	ARN	TRN	WCT	N	WRT	ARN	TRN	Start-P	End-P	BHRTN
1	5	6	0.83	1.25	3.56	2	2	1	1.41	1.99	300	514	1.71
2	8	9	0.89	1.12	2.73	1	1	1	1.02	1.02	360	520	1.44
3	5	5	1.00	1.17	2.07	1	2	0.5	1.01	1.02	517	588	1.14
4	2	2	1.00	1.37	1.88	1	1	1	1.21	1.21	478	502	1.05
5	4	6	0.67	1.18	2.46	1	1	1	1.02	1.02	759	1130	1.49
6	2	2	1.00	1.43	2.03	0	0	NA	1.00	1.00	67	381	5.69
7	5	5	1.00	1.18	2.31	1	3	0.33	0.99	0.95	385	440	1.14
8	2	2	1.00	1.16	1.34	0	0	NA	1.00	1.00	748	893	1.19
9	10	10	1.00	1.07	1.98	4	5	0.8	1.04	1.23	540	884	1.64
10	14	17	0.82	1.07	2.82	6	7	0.86	1.07	1.62	407	552	1.36
Ave.	5.7	6.4	0.921	1.199	2.318	1.7	2.2	0.811	1.079	1.207	456.1	640.4	1.785

Table 5. Performance results for Experiment 2-1

7.3.2 Experiment 2-2

In this experiment, the same data of 10 stock names of Experiment 2-1 were used. By concurrently monitoring the stocks of the 10 names for the in-sample period, just one trading rule was generated instead of one rule for each name (Generating a technical rule for each stock name and concurrently monitoring each stock with each corresponding rule in testing is possibly better, though. However, it was not examined in this study). Of course, the specified fitness was the same as that of Experiment 2-1: TRN. The results are shown in Table 6. The same experimental procedure was executed ten times. At the bottom of the table, the averages are shown. "NA" in the table suggests that the model trader looked for a trading signal but in vain. It is possibly showing overfitting. In Experiment 2-2, the trading opportunities (WCT and N) were increased in both of training and testing. Both of WRT and ARN of testing show still good values, 0.837 and 1.068. The average of the total returns (TRN) was also increased from 1.207 to 1.488 in Experiment 2-2. The average total return of 1.488 seems to be very good compared to that of Nikkei Average (Japanese representative stock index) that increased about 30% during the years 2003 to 2004. However, the average of buy and hold return of the ten stocks were about 1.785, far better than 1.488. So the result of the switching model is not surprisingly good. However, as mentioned previously, the return of 1.488 is better than the average return 1.351 which is without No. 6 stock (Sumitomo Heavy Industry) of which BHRTN is 5.69. This company did not give any opportunity of trading in the out-of-sample period of Experiment 2-1. It is probably because of too steep gain of the price. It is not clear, though, that same situation occurred in Experiment 2-2.

Experiment 2-2										
	In-Sample Period					Out-of-Sample Period				
Exp. No.	WCT	N	WRT	ARN	TRN	WCT	N	WRT	ARN	TRN
1	2	2	1.00	1.43	2.03	0	0	NA	1.00	1.00
2	10	10	1.00	1.15	3.94	4	5	0.80	1.09	1.51
3	2	2	1.00	1.43	2.03	0	0	NA	1.00	1.00
4	12	14	0.86	1.12	4.85	7	8	0.88	1.11	2.22
5	14	21	0.67	1.09	5.08	5	7	0.71	1.07	1.49
6	2	2	1.00	1.43	2.03	0	0	NA	1.00	1.00
7	10	10	1.00	1.15	4.14	5	5	1.00	1.14	1.96
8	19	22	0.86	1.10	6.95	4	6	0.67	1.05	1.34
9	8	11	0.73	1.12	3.10	5	5	1.00	1.13	1.86
10	10	10	1.00	1.15	3.94	4	5	0.80	1.09	1.50
Ave.	8.9	10	0.912	1.215	3.810	3.4	4	0.837	1.068	1.488

Table 6. Performance results for Experiment 2-2.

As seen in the table, when poor performances appeared for WCT in training (Exp. No. 1, 3, and 6), there were no trading in testing, TRN was just 1.00 in testing for all of the three. WCTs were all 2, very small, for the training of Exp. No.1, 3 and 6 in Table 6. It may be implying that in these simulations, the GA processes had fallen into local optimal points and that more optimization in training with more generations or larger population might bring about better performance for TRN in learning and testing.

7.4 Evaluation of stock switching

It is very difficult to evaluate the results of Experiment 2-2, the model trader of switching one stock to another. The average value of TRN should be considered very good, but not good enough compared with the average of buy and hold returns 1.785. However, it is true that the switching strategy brought about big improvement. That is, average return of Experiment 2-1 was 1.207, but that of Experiment 2-2 was 1.488, which is a note worthy improvement. This was brought about by switching stocks of the 10 names. Experiments with more stock names are strongly desired. It is also desired to examine how the model behaves when the market is in a down trend.

8. Third step experiments: Toward stable and profitable trading

What we have learned from the previous experiments of the two steps includes the following:

1. GA can find a good (may not a best) model trader.
2. By increasing the number of stock names, trading opportunity increases and it possibly makes overfitting less and trading profitable.
3. Stock switching is one of fairly good ways to handle multiple stock names.
4. Profitable trading might have relation with number of stock names.

The above prompted additional experiments with more stock names and the ones to seek quantity and quality of trading. The following experiments were designed:

8.1 Experiment 3-1

This experiment was done under the exactly same conditions as Experiment 2-2 except the number of stocks was increased to 50. The results are shown in Table 7. Performances were drastically improved from Experiment 2-2. The average TRN for the out-out-sample was increased from 1.488 to 2.20. Some TRNs are more than 3.0. The best is 5.77. This means the original amount became more than tripled or five times larger. This TRN was probably the result of stock switching of extremely good timing, which could hardly be expected for a small set of stock names. As we see the Exp. No. 7 in Table 7, the simultaneously obtained other indexes were also good. ARN was 1.293, which was extremely high; WRT was 1.00, perfect; WCT was 7, rather small.

One experiment (No. 6) resulted in loss, where TRN were 0.98. For this experiment ARN is 1.003, slightly profitable, though. This could happen. The results in general were very good, far better than the average return (1.49) of buy and hold for the 50 stocks. However, we don't know yet how to avoid the loss or how to pick up the one which would surely profitable or one of the best.

8.2 Experiment 3-2

Complex strategy often brings about poor results, overfitting. Even though good results are obtained for a learning period, the model traders are often inclined to pick up some particular profitable events, which would be hardly expected to happen again for an unknown future period, thus results for the unknown periods are frequently not good.

However, controlling fitness in a simple manner aiming to obtain good result may bring about good results. Since the quantity and quality are important for good profitable trading, profitable trading is considered to depend on the frequency of good quality trading. Since those two are considered to be essentially important in both periods of learning and testing, say in any situation, controlling quantity and quality of trading could create a good model trader. With this idea, we made the fitness of the GA to be Equation (10), which is composed of the following indexes. The similar idea was tried (Kurokawa, 2005). The indexes are the total return (TRN), the number of successful trading (WCT), the rate of profitable trading against the number of total trading (WRT) and the average return (ARN), all of which are essential factors for good quality trading. We consider those indexes are essentially related with profit (TRN) and consistently important in any situation. Another point of view for those indexes is the fitness has to be balanced among those indexes. Accordingly, the following index was tried for the fitness to control the experiment (Experiment 3-2).

$$fitness = TRN * WCT * WRT * ARN \tag{10}$$

The results are shown in Table 8. The average return (TRN) was improved from 2.20 to 2.32. One (No. 29) out of 30 experiments resulted in loss. The average successful number of trading (WCT) increased from 17.8 to 59.4. This is a sort of big change. The average return (ARN) changed from 1.059 to 1.008. This is also a significant change. The average winning rate (WRT) changed from 0.69 to 0.58, which is considered a slight change. The phenomena are understood as

1. The frequency of the trading increased significantly.
2. The individual trading profit rate decreased significantly.
3. The success rate decline was a little.

The above means that the obtained model trader would do frequent trading with each small profit. However, it is questionable if it is profitable after trading cost with such a low profit rate.

By the way, the company codes of the 50 stocks are as:

1515, 1803, 1812, 1928, 2002, 2503, 2897, 3402, 3407, 3861, 4005, 4183, 4452, 4502, 4519, 5001, 5333, 5401, 5405, 5706, 5802, 6113, 6301, 6448, 6702, 6753, 6902, 6991, 7011, 7203, 7733, 7735, 7912, 8001, 8015, 8031, 8233, 8332, 8355, 8604, 8801, 9007, 9064, 9101, 9104, 9301, 9302, 9501, 9502, 9719. As the codes suggest, the stocks are from almost all industries. All are from Tokyo Market of division I. The buy and hold rate for all the 50 stocks for the two year period 2003 to 2004 is 1.49. So the results of the above methods are said to be very good compared with B&H.

8.3 Experiment 3-3

In this experiment, simulation data period was shifted two years ahead. The learning period was from 2003 to 2004; the test period was 2005 to 2006. No other conditions were changed from Experiment 3-2. The results are shown in Table 9. The average return (TRN) was changed from 2.32 to 1.52, large decline, which is still considered to be good. This means performance could fluctuate somewhat extensively. No experiment out of the 30 was in loss. The average successful number of trading (WCT) changed from 59.4 to 48.8, not a drastic change. The average return (ARN) changed from 1.008 to 1.007, not significant. The average winning rate (WRT) changed from 0.58 to 0.63, which is also not significant. As a whole, trading characteristics were unchanged, though the total return shows a seemingly big change. By the way, the average B&H profit rate for the 50 stocks was 1.70, which is far better than the average performance of the experiment.

8.4 Experiment 3-4

With the results of Experiment 3-3, a question arose what would happen for Experiment 2-2 if the data period was changed. This experiment was for that. With the data period advanced two years with the previous 10 stock names and the fitness be just the TRN, unchanged, the experiments were done. The results were shown in Table 10. Seven out of 10 experiments had no trading for testing period. This is considered as that overfitting occurred or evolutions were immature. Any way the performance was very poor. TRN was just 1.09 compared with B&H was 1.63.

8.5 Experiment 3-5

This experiment was conducted with the fitness changed from Experiment 3-4, no other change. Fitness employed was Equation (10). The results are shown in Table 11.

The average TRN was 1.35 improved from 1.09; that for WCT was 16.5 from 2.2; that for WRT 0.70 from 0.87; and ARN 1.019 from 1.023. In general, the performances were similar to those of Experiment 3-3. This model trader is characterized as the sort of Experiment 3-3, frequent trading with each small profit.TRN was 1.35, which was still below B&H (1.63).

Experiment 3-1 50-Stocks Fitness=TRN										
	In-Sample(2001-2002					Out-of-Sample(2003-2004)				
Exp.No.	WCT	N	WRT	ARN	TRN	WCT	N	WRT	ARN	TRN
1	7	8	0.88	1.167	3.18	5	7	0.71	1.057	1.35
2	5	9	0.56	1.129	2.55	5	8	0.63	1.085	1.78
3	10	14	0.71	1.115	4.04	8	14	0.57	1.083	2.57
4	59	97	0.61	1.016	4.12	49	95	0.52	1.008	1.99
5	25	40	0.63	1.028	2.47	20	39	0.51	1.012	1.43
6	25	45	0.56	1.026	2.76	24	45	0.53	1.003	0.98
7	6	7	0.86	1.215	3.57	7	7	1.00	1.293	5.77
8	6	7	0.86	1.222	3.78	6	7	0.86	1.248	4.12
9	9	14	0.64	1.115	3.74	10	14	0.71	1.079	2.55
10	33	36	0.92	1.047	5.06	24	27	0.89	1.040	2.72
11	43	64	0.67	1.032	6.54	31	58	0.53	1.010	1.63
12	7	8	0.88	1.167	3.18	5	7	0.71	1.057	1.35
13	7	9	0.78	1.165	3.23	5	7	0.71	1.100	1.83
14	42	45	0.93	1.038	5.08	16	19	0.84	1.018	1.37
15	8	11	0.73	1.128	3.62	6	10	0.60	1.081	2.05
16	20	24	0.83	1.083	6.15	11	19	0.58	1.010	1.12
17	21	30	0.70	1.054	4.30	18	30	0.60	1.040	2.90
18	18	24	0.75	1.076	5.27	15	18	0.83	1.072	3.22
19	29	45	0.64	1.039	4.85	26	45	0.58	1.017	1.94
20	7	7	1.00	1.204	3.57	5	7	0.71	1.123	1.93
21	14	16	0.88	1.074	3.03	12	15	0.80	1.057	2.23
22	28	45	0.62	1.033	3.67	27	45	0.60	1.014	1.67
23	25	32	0.78	1.049	4.25	17	24	0.71	1.023	1.57
24	34	60	0.57	1.026	3.99	33	61	0.54	1.003	1.14
25	30	32	0.94	1.049	4.50	27	29	0.93	1.049	3.80
26	61	97	0.63	1.017	4.63	49	86	0.57	1.012	2.39
27	28	31	0.90	1.051	4.30	16	19	0.84	1.039	2.00
28	23	37	0.62	1.032	2.83	23	35	0.66	1.020	1.80
29	29	45	0.64	1.038	4.50	25	45	0.56	1.016	1.84
30	8	12	0.67	1.120	3.41	8	11	0.73	1.110	2.92
Ave.	22.2	31.7	0.75	1.085	4.00	17.8	28.4	0.69	1.059	2.20
						B&H Ave. Return Rate				1.49

Table 7. Performance results for Experiment 3-1.

Experiment 3-2 50-Stocks Fitness=TRN*WCT*WRT*ARN										
	In-Sample (2001-2002)					Out-of-Sample(2003-2004)				
Exp. No.	WCT	N	WRT	ARN	TRN	WCT	N	WRT	ARN	TRN
1	83	129	0.64	1.011	3.83	69	118	0.59	1.010	3.09
2	88	124	0.71	1.014	5.14	70	116	0.60	1.011	3.22
3	78	116	0.67	1.015	5.26	60	109	0.55	1.006	1.78
4	68	119	0.57	1.012	3.52	63	115	0.55	1.005	1.70
5	78	114	0.68	1.016	5.73	64	104	0.62	1.012	3.11
6	43	64	0.67	1.032	6.54	31	58	0.53	1.010	1.63
7	74	113	0.66	1.014	4.54	70	105	0.67	1.017	5.18
8	78	120	0.65	1.014	4.96	70	109	0.64	1.013	3.66
9	75	119	0.63	1.012	3.65	63	114	0.55	1.006	1.77
10	59	73	0.81	1.019	3.54	39	55	0.71	1.013	1.83
11	72	128	0.56	1.011	3.60	65	126	0.52	1.003	1.32
12	69	101	0.68	1.014	3.80	57	92	0.62	1.012	2.69
13	71	115	0.62	1.012	3.24	48	107	0.45	1.001	1.00
14	88	123	0.72	1.014	4.81	68	112	0.61	1.012	3.33
15	64	102	0.63	1.013	3.44	53	99	0.54	1.009	2.18
16	63	102	0.62	1.016	4.52	63	98	0.64	1.016	4.12
17	63	104	0.61	1.015	3.98	56	101	0.55	1.010	2.49
18	75	119	0.63	1.013	3.92	64	115	0.56	1.006	1.81
19	55	98	0.56	1.012	2.72	52	98	0.53	1.003	1.21
20	69	101	0.68	1.014	3.80	57	92	0.62	1.012	2.69
21	71	116	0.61	1.012	3.47	52	109	0.48	1.004	1.35
22	78	116	0.67	1.015	5.26	60	109	0.55	1.006	1.78
23	73	128	0.57	1.011	3.57	66	126	0.52	1.003	1.36
24	79	132	0.60	1.012	4.13	81	118	0.69	1.011	3.50
25	73	128	0.57	1.011	3.77	65	126	0.52	1.003	1.30
26	76	120	0.63	1.013	4.08	62	114	0.54	1.005	1.68
27	88	124	0.71	1.014	5.14	69	115	0.60	1.011	3.22
28	59	73	0.81	1.019	3.42	37	56	0.66	1.010	1.62
29	42	82	0.51	1.014	2.59	39	82	0.48	0.997	0.69
30	88	124	0.71	1.014	5.14	70	116	0.60	1.011	3.22
Ave.	71.3	110.9	0.65	1.014	4.17	59.4	103.8	0.58	1.008	2.32
						B&H Ave. Return Rate				1.49

Table 8. Performance results for Experiment 3-2.

Experiment 3-3 50-Stocks Fitness=TRN*WCT*WRT*ARN										
	In-Sample (2003-2004)					Out-of-Sample (2005-2006)				
Exp. No.	WCT	N	WRT	ARN	TRN	WCT	N	WRT	ARN	TRN
1	80	128	0.63	1.018	8.21	58	109	0.53	1.005	1.67
2	75	107	0.70	1.018	6.12	49	85	0.58	1.006	1.62
3	86	135	0.64	1.015	6.84	73	125	0.58	1.006	1.89
4	85	135	0.63	1.015	6.74	73	129	0.57	1.002	1.29
5	63	78	0.81	1.023	5.35	41	63	0.65	1.002	1.06
6	77	108	0.71	1.019	6.76	44	75	0.59	1.006	1.47
7	87	134	0.65	1.015	6.45	77	127	0.61	1.003	1.46
8	82	127	0.65	1.018	8.61	59	110	0.54	1.004	1.44
9	86	135	0.64	1.015	6.84	75	126	0.60	1.005	1.79
10	78	110	0.71	1.014	4.18	63	101	0.62	1.006	1.65
11	40	41	0.98	1.036	4.28	20	24	0.83	1.004	1.03
12	86	136	0.63	1.015	6.90	73	126	0.58	1.005	1.75
13	61	69	0.88	1.028	6.45	35	45	0.78	1.007	1.29
14	48	70	0.69	1.024	4.44	42	70	0.60	1.016	2.73
15	30	32	0.94	1.051	4.65	21	27	0.78	1.018	1.52
16	46	49	0.94	1.031	4.42	34	40	0.85	1.016	1.83
17	86	134	0.64	1.015	6.84	74	128	0.58	1.003	1.43
18	75	107	0.70	1.018	6.12	48	85	0.57	1.005	1.51
19	71	106	0.67	1.020	7.45	17	29	0.59	1.013	1.42
20	68	99	0.69	1.021	7.31	17	28	0.61	1.013	1.42
21	46	68	0.68	1.028	5.66	10	20	0.50	1.001	1.00
22	86	135	0.64	1.015	6.76	71	129	0.55	1.002	1.16
23	87	134	0.65	1.014	5.97	76	127	0.60	1.003	1.42
24	77	108	0.71	1.019	6.76	44	75	0.59	1.006	1.47
25	86	136	0.63	1.015	6.90	77	126	0.61	1.006	2.01
26	56	79	0.71	1.020	4.43	41	69	0.59	1.003	1.17
27	61	69	0.88	1.028	6.45	32	43	0.74	1.005	1.19
28	42	44	0.96	1.036	4.55	23	27	0.85	1.022	1.73
29	76	117	0.65	1.016	6.06	22	37	0.60	1.010	1.41
30	86	135	0.64	1.015	6.84	75	125	0.60	1.006	1.93
Ave.	70.4	102.2	0.72	1.021	6.18	48.8	81.0	0.63	1.007	1.52
						B&H Ave.Return Rate				1.70

Table 9. Performance results for Experiment 3-3.

Experiment 3-4 10-Stocks Fitness=TRN										
	In-Sample (2003-2004)					Out-of-Sample (2005-2006)				
Exp. No.	WCT	N	WRT	ARN	TRN	WCT	N	WRT	ARN	TRN
1	11	14	0.79	1.162	7.44	0	0	NA	1.000	1.00
2	15	19	0.79	1.127	8.81	0	0	NA	1.000	1.00
3	7	7	1.00	1.339	7.28	0	0	NA	1.000	1.00
4	9	10	0.90	1.230	7.31	0	0	NA	1.000	1.00
5	6	6	1.00	1.382	6.87	0	0	NA	1.000	1.00
6	9	9	1.00	1.274	8.35	0	0	NA	1.000	1.00
7	20	26	0.77	1.087	7.77	15	25	0.60	1.000	0.95
8	9	9	1.00	1.274	8.35	0	0	NA	1.000	1.00
9	13	15	0.87	1.149	7.49	2	2	1.00	1.111	1.23
10	20	26	0.77	1.091	7.99	5	5	1.00	1.116	1.71
Ave.	11.9	14.1	0.89	1.212	7.77	2.2	3.2	0.87	1.023	1.09
						B&H Ave. Return Rate				1.63

Table 10. Performance results for Experiment 3-4.

Experiment 3-5 10-Stocks Fitness=TRN*WCT*WRT*ARN										
	In-Sample (2003-2004)					Out-of-Sample (2005-2006)				
Exp. No.	WCT	N	WRT	ARN	TRN	WCT	N	WRT	ARN	TRN
1	41	43	0.95	1.031	3.73	28	33	0.85	1.017	1.68
2	33	33	1.00	1.040	3.62	0	0	NA	1.000	1.00
3	68	95	0.72	1.015	4.03	5	7	0.71	1.019	1.14
4	24	28	0.86	1.077	6.74	4	5	0.80	1.105	1.63
5	29	43	0.67	1.038	4.29	21	40	0.53	1.012	1.49
6	41	43	0.95	1.031	3.73	31	35	0.89	1.018	1.82
7	26	32	0.81	1.049	4.41	15	28	0.54	1.016	1.50
8	59	67	0.88	1.023	4.42	23	34	0.68	1.005	1.15
9	14	16	0.88	1.147	8.48	0	0	NA	1.000	1.00
10	51	68	0.75	1.017	2.86	38	59	0.64	1.002	1.05
Ave.	38.6	46.8	0.85	1.047	4.63	16.5	24.1	0.70	1.019	1.35
						B&H Ave. Return Rate				1.63

Table 11. Performance results for Experiment 3-5.

8.6 Discussions for third step experiments

There are three points to discuss for this section. First point is that some model traders demonstrated extremely good performances. This happened for a somewhat large group of stock names, 50. The results so far shown are good from a general point view as well. However, it does not mean we can always get such good results. Nonetheless, it demonstrated that some model traders exist which could bring about extremely good results maybe depending on the situations and timing. Therefore we should keep investigating what kinds of stock groups are profitable.

Second is that the same GA process works differently depending on the time period. One time, it works extremely well; on another time it works very poorly. We like to seek what would cause such differences. This may be significant because some solutions may exist behind the difference.

Third is a special point of the second. It is related with what is essentially important for good model traders for both of training period and testing one, say in any situation. It is probably true that there is not such a consistent general matter. However, the two experiments Experiment 3-2 and Experiment 3-4 demonstrated that the fitness of Equation (10) brought the change to the trading quantity and quality as well as the TRN increase. That is, the winning number of trading increased and the individual trading profit rate declined. This suggests that it may be possible to control the quantity and quality of trading by the fitness and that there is also a possibility of profit related general matters around those indexes. It is worth investigating those indexes further. We hope there is even only a bit of essential generality around them.

9. Discussions for what to seek

Much difficulty comes from the facts that the spaces of in-sample and out-of-sample data are so wide that cause and effect relations are hard to find. In addition, there may exist many factors which do not appear in data. The space widths, the ever-changing market conditions and often contradicting data of the market are considered to be major causes of overfitting and instability.

What we need is a stable good performance model trader with generality. It is, however, true that we can find a good trader for in-sample data. We want the created trader will work well again for unseen out-of-sample data with good probability. What we like to seek is the factors, which are extracted for the in-sample data, which play important roles for the out-of-sample data, too. Those factors should play important roles for out-of-sample data thus reducing overfitting. One of the important points is that frequently repeated events for in-sample data would happen again, hopefully repeatedly and frequently for unseen out-of-sample data.

In our social world, events which happened many time for a period are expected to happen many times again for other period, too. Stock market events are not exceptions. So it is considered important to find a model trader that can catch the profitable and frequent trading events.

In the above sense, the factors of Equation (10) are considered important. The equation for the fitness affected significantly the quantity and quality of trading. Those of profit related indexes are considered consistently important for both of in-sample and out-of-sample periods and in any situation, should play important roles for organizing a good stable model trader even if it may depends on random processes. Including those indexes, we should keep seeking indexes which are of more consistent, general and profit bringing matters.

10. Conclusion

The following are concluded:

1. Overfitting is defined somewhat formally as Equation (5).
2. Trading opportunity increased by increasing data size. This decreased overfitting. Some experiments verified it numerically.
3. Trading chances appear in many places among stock names.
4. Stock switching demonstrated good results.
5. Increased number for stock names demonstrated some extremely good results.
   Some model traders with switching from one stock to another showed the big performance improvement in experiments.
6. The same GA process which worked well in some situations does not necessarily work well for other situations.
7. Generally acceptable profit related matters which work well consistently in any situations should be sought.
8. Equation (10) for fitness demonstrated the potential ability to create a stable profitable model trader, and experiments suggests that equation and the related indexes have some potential to control quantity and quality of trading.

11. References

Becker, L. A. & Seshadri, M. (2003a). Comprehensibility and Overfitting Avoidance in GENETIC PROGRAMMING, *Worcester Polytechnic Institute, Computer Science Technical Report WPI-CS-TR-03-09*, Retrieved from http://www.cs.wpi.edu/Resources/techreports.html?/

Becker, L. A. & Seshadri, M. (2003b). GP-Evolved Technical Trading Rules can Outperform Buy and Hold, *Worcester Polytechnic Institute, Computer Science technical Report WPI-CS-TR-03-16*, Retrieved from http://www.cs.wpi.edu/Resources/techreports.html?/

Geurts, P. (2005). Bias vs Variance Decomposition for Regression and Classification In: *The Data Mining and Knowledge Discovery Handbook*, Maimon, O. & Rokach, L. (Eds.), pp.749-763, Springer, ISBN-10: 0-387-24435-2

He, H.; Chen, J.; Jin, H. & Chen, S. H. (2007). Trading Strategies Based on K-means Clustering and Regression Models, In *Computational Intelligence in Economics and Finance, Vol. II*, Chen, S. H. ; Wang, P. P. & Kuo, T. W. (Eds.), pp.123-134, Springer

Kurokawa, T. (2005). Evolutionary Stock Trading Method by Effective Catching of Market Recovery Reaction, *The Proceedings of the 6th International Symposium on Advanced Intelligent Systems*, pp.806-811, Yeosu, Korea

Kurokawa, T. (2007). Evolutionary Method to Optimize Composite Indicator for Market Timing, *The Proceedings of the 8th International Symposium on Advanced Intelligent Systems*, pp.264-269

Kurokawa, T. (2008). On Overfitting of Technical Market Timing with Evolutionary Process—Effects of In-Sample Data Selection, *The Proceedings of the 9th APIEMS-2008,* pp.451-458, Bali, Indonesia

Kurokawa, T. (2009). Learning and Technical Market—Effects of In-Sample Data Selection, *Journal of Advanced Computational Intelligence and Intelligent Informatics*, Vol.13, No. 6, pp.726-730, ISSN 1343-0130

Kurokawa, T. (2011). Stock Trading with Genetic Algorithm-Switching from One Stock to Another, *Journal of Communication and Computer*, Vol.8, No.2, pp.143-149, ISSN 1548-7709

Lam, S. S. ; Lam, K. P. & Ng, H. S. (2002). Genetic Fuzzy Expert Trading System for NASDAQ Stock Market Timing, In: *Genetic Algorithms and Genetic Programming in Computational Finance*, Chen, S. H., (Ed.), pp.197-217, Kluwer Academic Publishers, ISBN 0-7923-7601-3, Massachusetts, USA

Lin, L. ; Cao, L. & Zhang, C. (2005). Genetic Algorithms for Robust Optimization in Financial Applications, *Proceedings of the Fourth IASTED International Conference of Computational Intelligence*, pp.387-391

Lohpetch, D. & Corne, D. (2010). Outperforming Buy and Hold with Evolved Technical Trading: Daily, Weekly and Monthly Trading In: *Application of Evolutionary Computation 2010 Proceeding, Part II*, Chio, C. D. et al. (Eds.) , pp. 171-181, Springer-Verlag, Berlin Heidelberg, ISBN-10: 3-642-12241-8

Mabu, S. ; Hirasawa, K. & Furuzuki, T. (2007). Trading Rules on Stock Markets Using Genetic Network Programming with Reinforcement Learning and Importance Index, *Transaction of IEE Japan C,* Vol.127, No. 7, pp.1061-1067

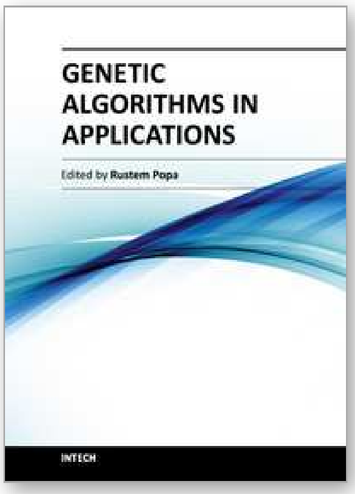
Neely, C. ; Weller, P. & Ditmar, R. (1997). Is Technical Analysis In The Foreign Exchange Market Profitable? --- A Genetic Programming Approach, *Journal of Financial and Quantitative Analysis,* Vol.32, No.4, pp405-427

Pavlidis, N. G. ; Pavlidis, E. G. ; Epitropakis, M. G. ; Plagianakos, V. P. & Vrahatis, M. N. (2007). Computational Intelligence Algorithms for Risk-Adjusted Trading Strategies, *IEEE Congress on Evolutionary Computation (CEC),* pp.540-547, ISBN 978-1-4244-1339-3

Potvin, J. Y.; Soriano, P. & Vallee, M. (2004). Generating Trading Rules On The Stock Markets With Genetic Programming, *Computers & Operations Research*, Vol. 31, pp.1033-1047

Takizawa, T. (1999). *SP-Wave Method*, Pan-Rolling Publishing, Tokyo, Japan

Wang, J. & Chen, S. (1998). Evolutionary Stock Trading Decision Support System Using Sliding Window, *Proceedings of the IEEE Conference on Evolutionary Computation*, pp.253-258



Genetic Algorithms in Applications

Edited by Dr. Rustem Popa

ISBN 978-953-51-0400-1
Hard cover, 328 pages
Publisher InTech
Published online 21, March, 2012
Published in print edition March, 2012

Genetic Algorithms (GAs) are one of several techniques in the family of Evolutionary Algorithms - algorithms that search for solutions to optimization problems by "evolving" better and better solutions. Genetic Algorithms have been applied in science, engineering, business and social sciences. This book consists of 16 chapters organized into five sections. The first section deals with some applications in automatic control, the second section contains several applications in scheduling of resources, and the third section introduces some applications in electrical and electronics engineering. The next section illustrates some examples of character recognition and multi-criteria classification, and the last one deals with trading systems. These evolutionary techniques may be useful to engineers and scientists in various fields of specialization, who need some optimization techniques in their work and who may be using Genetic Algorithms in their applications for the first time. These applications may be useful to many other people who are getting familiar with the subject of Genetic Algorithms.