Genetic Commerce – Intelligent Share Trading

Clive Vassell

Harrow Business School,
University of Westminster,
London
Email: vasselc@wmin.ac.uk,
URL: users.wmin.ac.uk/~vasselc

Abstract. In time, it seems feasible that genetic algorithms will help to achieve similar levels of productivity gains in many service domains as has been achieved in line and, more recently, batch manufacture. And e-commerce will embody the standards and guidelines necessary to enable managers to make the most of the possibilities offered by this development. It will help them to both satisfy and retain the custom of their clients; and make it possible to operate in a highly efficient and effective manner. The paper discusses the nature of these changes and it assesses some of their implications for organisations and their management. It introduces the concept of intelligent share trading; a future manifestation of these developments. And it talks about the important role artificial intelligence, particularly genetic algorithms, will play in these systems.

Electronic Commerce

The Internet looks set to become a key plank in the infrastructure needed to support a new era of business development. The promise of a network connecting all significant economic agents of the world (both human and software) and many of the devices on which they rely creates the possibility of a huge array of information services [1].

If the Internet is, in large measure, the infrastructure which facilitates this new era, e-commerce is the 'business protocol' which determines the standards and norms by which trade is conducted in this new context.

It covers such issues as electronic data interchange (EDI) between the various members of the virtual supply chain, the payment systems which are to be used and/or permitted, and the maintenance of the levels of security necessary to reassure customers and potential customers.

Just as importantly, it encapsulates the understanding gleamed about what works and what doesn't in the information age; the 'strategic protocol' which leads to satisfied and loyal customers as well as robust and profitable businesses.

It is, as yet, early days. We still have a considerable amount to learn; and we still have many tried and trusted approaches carried over from the preceding era which will need to be unlearned as they no longer apply.

A few indicators of the likely form of these protocols are beginning to emerge however. First of all, the customer is going to be better placed tomorrow than he or she is today [2]. The information age will make information available to the connected world in abundance; it will no longer be the preserve of the resourceful or powerful.

Information on available products, prices, relative performance, cost of production, methods of production and distribution, environmental friendliness, suitability for various tasks and/or users, and much, much more will be available to all those interested to look for it [3].

And looking for such information will become progressively easier. We suffer from information overload now not so much because there is too much information out there but rather because the information searching and filtering tools we have at present are not sufficiently sophisticated.

As these tools improve, so our ability to more effectively manage large amounts of information will increase as will our ability to be selective about the theme of the information presented, its format, and its level of detail.

It will become successively more difficult, indeed counterproductive, for large suppliers to misrepresent their products or services. Should they do so, an army of empowered consumers may abandon their offerings and might well petition (via the Net) industry watchdogs, MPs, consumer groups or TV programs or web sites, or any others they feel appropriate to take action against the offending firm.

Furthermore, for the foreseeable future there will remain a need for effective logistics [4]. Indeed e-commerce is likely to increase the need for this and yet it is an area that is often overlooked. Even information services require facilitating goods (equipment and consumables) and the more dispersed the service provision, the more carefully the supporting logistics will have to be planned and implemented.

Recently, data mining has become an area of considerable interest to organisations. Many large firms have huge data warehouses full of transaction data, management information, information on the external environment, and information on much else of potential value. Data mining approaches, including artificial intelligence, can be very useful in making sense of this mountain of data and using that understanding to improve decision making [5], [6], [7].

Artificial intelligence is being used in a wide range of applications. It is being used to better facilitate manufacturing [8], and to make intelligent agents more effective [9], and in a host of other applications and domains in between.

And, inevitably, it is being used in the financial arena. Neural networks and genetic algorithms are the preferred AI environments in this area. The vast amounts of data available and the complexity of the applications make them particularly well suited to the domain. And thus they are being used for a range of financial applications, including stock market prediction [10], [11].

These are but a few examples of the kind of insight of this new order which e-commerce will have to encapsulate. There will be many more which will emerge over the years ahead.

Organisations will have to be sensitive to the effectiveness of the initiatives they introduce and be ready to respond rapidly where and when required. This will be no easy task but if it is done well enough and soon enough, the rewards could be a place among the great and the good of the new order.

Intelligent Share Trading

So how might all this work in practice? Well, one important application of e-commerce is share trading. There are many organisations which provide a service that enables customers to buy and sell shares of their choice on one or more stock exchange(s).

Typically the user is required to select the shares he or she wishes to buy or sell and carry out the transaction. The system makes it possible for the user to do so simply and quickly, and typically provides up-to-date prices and some relevant news. But it does not make the choices and it does not automatically conduct transactions.

In principle, however, it would be perfectly feasible for such systems to be extended to include these two additional functions. The systems could use a selection strategy to choose which shares to buy and a deselection strategy to determine which ones to sell. It could then conduct the buy and sell transactions whenever the appropriate conditions applied.

The selection and deselection strategies could be specified by the user so that the system behaves in the way that the user would but does so automatically and responds to changed circumstances almost instantaneously. Alternatively, the systems could use artificial intelligence to determine appropriate selection and deselection criteria.

The latter is, in essence, an intelligent share trading system. It would automatically trade on its user's behalf according to criteria it has determined; the user simply needs to set the broad performance goals (such as strong growth or modest volatility), and periodically review the performance of the portfolio and the history of transactions.

The Constituents of an Intelligent Share Trading System

The main constituents of an intelligent share trading system would be the data collection module, the share trading module, the share selection module and the strategy optimisation module.

The data collection module would collect share price information and store it in a database ready for processing by the share selection and strategy optimisation mod-

ules. The selection module would apply the current selection and deselection strategy to determine whether any shares ought to be bought or sold at present. If so it would request that the share trading module conduct a trade. The trading module would then buy or sell the requested quantity of the required share.

The strategy optimisation module would run periodically and, using historical data, determine which investment/trading strategy would have given optimal results – in line with the broad performance objectives specified by the user. This would then become the strategy to be applied by the share selection module.

Genetic algorithms would be used to effect the optimisation. They are arguably one of the best tools for finding optimal or near optimal solutions given an infinite or very large range of potential solutions.

In fact the genetic algorithms would operate on two levels. On one level they would find the optimum strategy given a set of attributes; on the second level they would find the set of attributes which yield the best strategy given a larger universe of attributes.

(This would help to ensure that the system comes up with both the optimal set of share attributes on which to base many of the better strategies, and the optimum strategy itself).

The system would be designed to devise strategies which were robust or, in other words, optimised for say ten years (rather than 'over optimised' for a short period of time and/or a particular set of market circumstances). This should mean that the strategy selected would not need to change often unless the user changed the overall performance objectives.

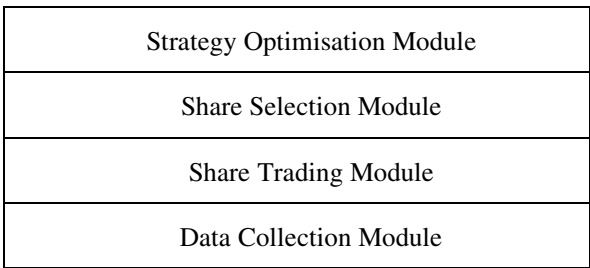| Strategy Optimisation Module |
| Share Selection Module |
| Share Trading Module |
| Data Collection Module |

Fig. 1. The main components of an intelligent trading system

Genetic Algorithms and Optimisation

Genetic algorithms apply the principles of natural selection to finding optimum solutions to a problem. They operate on a population of potential solutions and apply the principles of selection, crossover and mutation to produce a new generation of candidate solutions.

The selection operation is used to choose the best candidates from the population by testing each solution against a fitness or target function. The crossover operator is used to produce a child solution from two parent solutions by combining elements of the chromosomes of one parent with elements of the chromosomes of the other. The mutation operator introduces an element of randomness into each population.

The combination of these three operators leads to new generations of solutions which tend to improve in performance in relation to previous generations but which will not simply converge on sub-optimal solutions, in much the same way as they help living organisms to thrive.

The Nature of the Optimisation

So what might the optimisation dialogue look like? It would have screen displays similar to some of the screens taken from one of the models I am currently using for my research in this area. I shall use this model to give an example of how this part of the intelligent share trading system might look and feel.

An important component of the Strategy Optimisation Module of the intelligent share trading system would be the performance summary screen. An example of how it might look is shown below.

The summary would indicate what the current share selection strategy was and how that strategy would have performed in the past. In fig. 2 no strategy has been selected. This is indicated by the fact that the chromosome values are all zero.

Fig. 2. The strategy performance summary screen

		3mCnt	3mMean	3mMin	6mCnt	6mMean	6mMin	1yCnt	1yMean	1yMin
Summary										
Year 1		8	-8.76	-25.60	8	-5.46	-23.10	8	2.45	-32.40
Year 2		17	2.54	-8.00	17	8.88	-40.60	17	29.51	-40.60
Year 3		26	7.38	-23.90	26	36.08	-26.40	26	25.37	-54.00
Year 4		23	5.87	-17.90	23	13.98	-28.40	23	11.90	-47.30
Year 5		101	-0.45	-50.40	101	-6.04	-65.80	101	-1.36	-80.80
Year 6		46	-3.43	-37.50	46	-5.02	-38.20	46	-10.81	-72.22
In Sample Min		8	-8.76	-50.40	8	-6.04	-65.80	8	-1.36	-80.80
In Sample Median		26	-0.45	-25.60	26	-5.46	-26.40	26	2.45	-54.00
In Sample Mean		45	-0.61	-33.30	45	8.19	-38.43	45	8.82	-55.73
Out of Sample Min		17	-3.43	-37.50	17	-5.02	-40.60	17	-10.81	-72.22
Out of Sample Median		23	2.54	-17.90	23	8.88	-38.20	23	11.90	-47.30
Out of Sample Mean		29	1.66	-21.13	29	5.95	-35.73	29	10.20	-53.37
Overall Min		8	-8.76	-50.40	8	-6.04	-65.80	8	-10.81	-80.80
Overall Median		25	1.05	-24.75	25	1.93	-33.30	25	7.18	-50.65
Overall Mean		37	0.52	-27.22	37	7.07	-37.08	37	9.51	-54.55
Chromosomes										
Previous 3 Month Price Rise	Cppr3m	0	0-100							
Previous 6 Month Price Rise	Cppr6m	0	0-100							
Previous 1 Year Price Rise	Cppr1y	10	0-100							
Price/Earnings Ratio	Cpe	0	0-100							
Dividend Yield	Cyld	5	0-100							
Estimated Capital (£100k)	Ccap*100	0	0-100							
1 Year Earnings Per Share Growth %	Cepsgr	0	0-100							
1 Year Profit Growth %	Cprofgr	0	0-100							
1 Year Dividend Growth %	Cdivgr	0	0-100							
1 Year Turnover Growth %	Ctogr	0	0-100							

Fig. 3. The strategy performance summary screen with a strategy selected

In fig. 3, the selection strategy is:

> Select those shares where the share price has risen by 10% or more in the previous year and the previous year's dividend has been 5% or more of the (prevailing) share price.

It is important to understand that this selection strategy would have been applied at the start of each of the six years represented in the data set. And the size and performance of the resulting share portfolio for each of the six years for the subsequent three months, six months and one year are recorded at the top of the screen.

The three 'in sample' rows provide summaries of the in sample data (in this case years one, three and five). The three 'out of sample' rows provide summaries of the out of sample data (years two, four and six) and the three overall summary rows provide summaries of all the data (years one to six).

The target cell is also shown (the in sample minimum of the one year means). The target function is to maximise the value of this cell providing the associated count (the in sample minimum of the counts of the shares selected for each year) is at least ten.

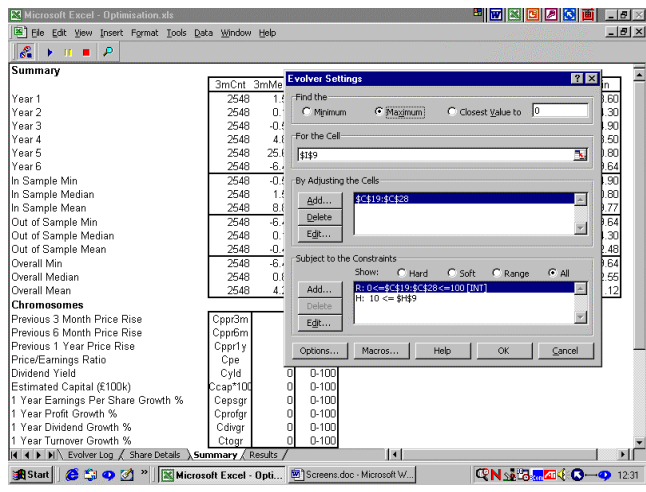The target function is defined in the screen below:

Fig. 4. Defining the *target function*

The Test attribute (or field) is used to determine which shares in the database meet the current share selection criteria. The contents of the Test attribute can perhaps best be explained if it is laid out in a logical format. This is done below. Here we can see how the genetic algorithm operates at the two levels spoken about earlier.

The Contents of the Test Attribute

```
=And (
Or (Cppr3m=0,    And (PriceRisePrev3m<>"",    PriceRisePrev3m>=Cppr3m) ),
Or (Cppr6m=0,    And (PriceRisePrev6m<>"",    PriceRisePrev6m>=Cppr6m) ),
Or (Cppr1y=0,    And (PriceRisePrev1y<>"",    PriceRisePrev1y>=Cppr1y) ),
Or (Cpe=0,       And (PERatio<>"",            PERatio<=Cpe) ),
Or (Cyld=0,      And (DividendYield<>"",      DividendYield>=Cyld) ),
Or (Ccap=0,      And (EstCapital<>"",         EstCapital>=Ccap) ),
Or (Cepsgr=0,    And (EPSgrowth<>"",          EPSgrowth>=Cepsgr) ),
Or (Cprofgr=0,   And (Profitgrowth<>"",       Profitgrowth>=Cprofgr) ),
Or (Cdivgr=0,    And (Dividendgrowth<>"",     Dividendgrowth>=Cdivgr) ),
Or (Ctogr=0,     And (Turnovergrowth<>"",     Turnovergrowth>=Ctogr) )
)
```

Where a chromosome has a value of zero, that chromosome plays no part in selecting shares. It is, in effect, switched off. Where it has a non-zero value, it is used in the selection process. Thus the genetic algorithm will, at any one time, be selecting which chromosomes are to be part of the selection criteria and which are not.

Where a chromosome has a non-zero value, the Test attribute selects those shares which have both a non-blank value in the appropriate attribute and the value of the attribute meets the criteria level of the associated chromosome.

In principle, this means we can have a large number of chromosomes but that, at any moment in time, the genetic algorithm will generally be using a few of them only. It means we should be able to test the effectiveness of a wide range of chromosome permutations while at the same time testing a range of selection strategies using each of those permutations. And we should be able to do this without having to build a whole suite of models.

(I think this principle may well reduce the likelihood of 'overfitting' in a relatively straightforward manner. Whether this particular application of the principle is likely to work remains to be seen.)

Conclusions

It should be noted that my main concern when trying to identify suitable strategies is robustness. In other words, I am interested in strategies which give good returns year in, year out rather than strategies which give high mean returns but are accompanied by high levels of volatility.

This is why my target function is the minimum one year mean (rather than the mean of all years). I am particularly interested in strategies which provide positive returns in all test years (both in sample and out of sample).

Perhaps in part because of the quite specific nature of my requirements, the results to date have not been entirely encouraging. I have not, so far, come across any strategy which meets these requirements (though there are one or two which show modest losses in one year only).

However the work continues. I plan to extend the range of candidate chromosomes and possibly introduce both maximum value and minimum value chromosomes to see whether this improves the performance of the best strategies.

While the results of the exercise are not yet terribly encouraging, there is some evidence to suggest that the search should prove fruitful in the end [12] and there is a prominent theoretical framework which is compatible with this kind of expectation [13], [14], [15].

Wider Implications

We are entering a new era in the history of commercial activity. The understanding which helps to crystallise this new era will be the body of knowledge which e-commerce will comprise. It will be the content of the MBAs of the online business schools of tomorrow and will enable executives to profitably steer their firms in the decades ahead.

The real winners however will probably not have learned much of what is critically important about e-commerce from business schools but rather by being the first to live them and learn from them in their own organisations [16]. And by being sufficiently capable as managers to capitalise extensively on the lead they so gain.

It is organisations like these that will fashion the new era; and it is those who study these organisations that will identify the e-commerce protocols associated with this novel form.

Intelligent share trading is an example of this kind of development. It results from a fusion of electronic commerce, artificial intelligence and online trading.

These systems are likely to prove profitable for their suppliers and their users alike. And the organisations who are the first to introduce robust and user friendly examples will in all probability do very well indeed.

And these systems will inevitably make extensive use of artificial intelligence. The huge amounts of historical data which can be called upon to facilitate understanding and the complexity of the application areas will make the use of data mining techniques and tools very valuable. And neural networks and genetic algorithms will likely prove extremely pertinent.

And genetic algorithms (and possibly hybrid approaches) will be of particular value in situations where it is necessary to both carry out appropriate actions on behalf of the users and explain to the users the underlying strategy behind those actions.

So important might this kind of artificial intelligence become that perhaps genetic commerce will be the most appropriate way to describe systems of the type outlined in this paper. Indeed its popularity might well signal the next phase of the incessant rise of the machine.

References

1. Aldrich Douglas F, 'Mastering the Digital Marketplace: Practical strategies for competitiveness in the new economy', John Wiley, 1999
2. Hagel John, Armstrong Arthur G, 'Net Gain: Expanding markets through virtual communities', Harvard Business School Press, 1997
3. Evans Philip, Wurster Thomas S, 'Getting Real About Virtual Commerce', Harvard Business Review, November-December 1999
4. Jones Dennis H, 'The New Logistics: Shaping the new economy', Ed: Don Tapscott, Blueprint to the Digital Economy: Creating wealth in the era of e-business, McGraw Hill, 1998
5. Gargano Michael L, Raggad Bel G, 'Data Mining – A Powerful Information Creating Tool', OCLC Systems & Services, Volume 15, Number 2, 1999
6. Lee Sang Jun, Siau Keng, 'A Review of Data Mining Techniques', Industrial management and Data Systems, Volume 101, Number 1, 2001
7. Bose Indranil, Mahapatra Radha K, 'Business Data Mining – A Machine Learning Perspective', Information & Management, Volume 39, Issue 3, December 2001
8. Burns Roland, 'Intelligent Manufacturing', Aircraft Engineering and Aerospace Technology: An International Journal, Volume 69, Number 5, 1997
9. Saci Emilie A, Cherruault Yves, The genicAgent: A Hybrid Approach for Multi-Agent Problem Solving', Kybernetes: The International Journal of Systems & Cybernetics, Volume 30, Number 1, 2001

10. Wittkemper Hans-Georg, Steiner Manfred, 'Using Neural Networks to Forecast the Systemic Risks of Stocks', European Journal of Operational Research, Volume 90, Issue 3, May 1996

11. Back Barbo, Laitinen Teija, Sere Kaisa, 'Neural Networks and Genetic Algorithms for Bankruptcy Predictions', Expert Systems With Applications, Volume 11, Issue 4, 1996

12. Bauer Richard J, 'Genetic Algorithms and Investment Strategies', John Wiley & Sons, February 1994

13. Fama Eugene F, French Kenneth R, 'The Cross-Section of Expected Stock Returns', The Journal of Finance, Volume 47, Number 2, June 1992

14. Fama Eugene F, French Kenneth R, 'Size and Book-to-Market Factors in Earnings and Returns', The Journal of Finance, Volume 50, Number 1, March 1995

15. Fama Eugene F, French Kenneth R, 'Multifactor Explanations of Asset Pricing Anomalies', The Journal of Finance, Volume 51, Number 1, March 1996

16. Senge Peter M, 'The Fifth Discipline: The art and practice of the learning organisation', Business (Century/Arrow), 1993