Natural Computing and Finance

Anthony Brabazon
&
Michael O'Neill

PPSN 2010
12 September 2010

www.fmc-cluster.org

UCD Natural Computing Research & Applications Group (NCRA)
http://ncra.ucd.ie

Overview

- Introduction
- Biologically-inspired Algorithms
- A Tour of Some Financial Applications …
  - Optimisation
  - Model-induction
  - Agent-based Modelling

Research Group

- Natural Computing Research and Applications Group (http://ncra.ucd.ie)
  - Interdisciplinary research group based in UCD's CASL Institute
  - (http://casl.ucd.ie)
- Development of NC algorithms / theory
- Application areas
  - Finance, bioinformatics, architecture, engineering, sound synthesis, game AI etc. etc.
- Staffing
  - Five faculty
  - 15 PhDs
  - Five Post docs

Financial Mathematics and Computation Cluster
– SFI Strategic Research Cluster

FMC² Financial Mathematics and Computation Research Cluster

Core Research Team
6 Lead academics (3 universities)
15 academic collaborators (from 5 other universities)
3 Post Doctoral Researchers
15 PhD Researchers
5 Industry Partners

Initial Research Activities
- Robust Asset Allocation
- Fund Performance Evaluation
- Crashes and Portfolio Choice
- Information Theory and Financial Markets - Model Selection and Complexity
- Grammatical Evolution for Asset Allocation – equity and fixed income
- Algorithmic Trading
- Asset Pricing and Risk
- Risk Management of Real Estate
- Pension Risk
- Time-series Dynamics of Multivariate Return Distributions
- Semi-parametric Estimation of Portfolio Risk
- Copulas, Fractals and Chaos

sfi science foundation ireland

Natural Computing

Inspire Design of *
Natural Processes as Computational Machines
Understanding Natural Processes

- *Working definition for this presentation …*
  - "the development of computational algorithms using metaphorical inspiration from systems and phenomena that occur in the natural world"
- Many of the best known algorithms in have been derived from a biological or social inspiration
- What makes the biological metaphor interesting when designing artificial problem solvers?

Biologically-inspired Algorithms

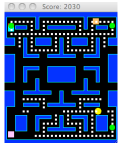
- Biological organisms earn a living in 'difficult' environments
  - Typically "high-dimensional" and dynamic
- Mechanisms have arisen which assist the 'survivability' / adaptability of populations of biological creatures in these environments
- These are potentially useful in helping inspire us when designing algorithms to attack interesting real-world problems in the finance (and other) domain(s)

Dynamic Environments

- Biological responses to dynamic environments include …
  - Populations rather than individuals
    - Multiple individuals generates diversity
    - Multiple 'probes' (learning trials) of the environment
    - Lower risk of extinction if you have diversity
  - Multiple mechanisms for learning
    - Personal lifetime learning capability
    - Social learning (communication between individuals)
    - Genetic learning
  - Mechanisms for maintaining a memory of good past solutions
  - Mechanisms for generating diverse new individuals (solutions)
  - Fitness-based (de)selection
  - Focus is on survival not optimisation (i.e. robustness)

Earning a Living …

- Let's take a simple example of the problem of 'earning a living' in a dynamic environment where the future actions of other agents are unknown …



Earning a Living …

- At each time step in the program one of nine rules (in decreasing priority) is fired

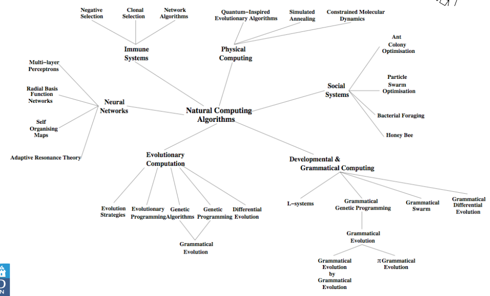
#Rule 1:

IF (distance(nearest_power_pill) ≤ 5(3*)) AND (4 ≤ distance(nearest_ghost) ≤ 8) AND (distance(ghost_nearest_to_the_nearest_power_pill) ≥ 6(4*)),

THEN stop moving and ambush (enter the ambush state) at the corner or cross point near the nearest power pill waiting for a ghost to come closer, where distance (nearest_power_pill) is the distance from Ms. Pac-Man to the nearest power pill, distance(nearest_ghost) the distance from Ms. Pac-Man to the nearest ghost, and distance(ghost_nearest_to_the_nearest_power_pill) the distance from Ms. Pac-Man to the ghost nearest to the power pill nearest to Ms. Pac-Man, and the numbers with * in the parentheses are those for the second stage of the game.

- How can we find good 'rules' for surviving in this (or any other…) environment?

Natural Computing Algorithms



Financial Application Areas of NC

i. Optimisation

ii. Model induction

iii. Agent-based modelling

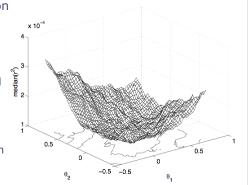
Survey Articles

Brabazon, A., O'Neill, M. and Dempsey, I. (2008). An Introduction to Evolutionary Computation in Finance, IEEE Computational Intelligence Magazine, 3(4):42-55.

Brabazon, A., Dang, J., Dempsey, I. and O'Neill, M. (2010). Natural Computing in Finance: A Review, in Handbook of Natural Computing: Theory, Experiments and Applications G. Rozenberg, T. Baeck and J. Kok (eds),), Berlin: Springer (forthcoming in December 2010).

Financial Application Areas of NC

- Optimisation
  - GA, DE, ACO, PSO … (we will focus on GA)
  - Heuristic methods can be useful particularly,
    - In problems which traditional optimisers find difficult (multiple local optima etc.)
    - In facilitating robust estimation of parameters (e.g. extreme values can make MSE-based parameters unstable) – removing the 'closed form solution' requirement
  - Primary applications in finance are parameter estimation / model calibration

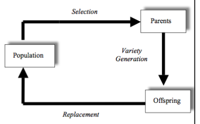


Exemplar mapping from coefficients into median squared residual illustrating multiple local optima (LMS estimation) (Gilli and Schumann, 2009)

Evolutionary Computation

Canonical evolutionary algorithm

Initialise population
WHILE (Termination condition False)
    Calculate fitness of each individual
    Select parents
    Create offspring
    Update population
ENDWHILE



$$x[t+1] = r(v(s(x[t])))$$

A Simple Optimisation Example

- Sometimes designing the genotype to phenotype mapping is simple
  - For example, suppose we want to design a genotype to encode three coefficients for a linear regression model of the form …

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

  - The genotype could be a real-valued string, encoding the three model coefficients

-3.1245	5.6219	11.3411

A Simple Optimisation Example

	β_0	β_1	β_2		Fitness (MSE)
1	-3.1245	5.6219	11.3411		0.3245
2	-4.5612	-0.2317	6.1311		0.7436
3	2.3412	1.6432	2.7811		0.6718
…	…	…	…		…
n	-3.6245	4.8219	13.3411		0.3015

A Simple Optimisation Example

	β_0	β_1	β_2		Fitness (MSE)
1	-3.1245	5.6219	11.3411		0.3245
2	-4.5612	-0.2317	6.1311		0.7436
3	2.3412	1.6432	2.7811		0.6718
…	…	…	…		…
n	-3.6245	4.8219	13.3411		0.3015

Suppose these two 'good' parents are chosen

A Simple Optimisation Example

	β_0	β_1	β_2		Fitness (MSE)
Child 1	-3.3745	5.2219	12.3411		…

We generate a 'child' solution by applying a pseudo-crossover operation to the two parents

Crossover uses information from both parents (recombines their good information – here using a simple averaging process)

A Simple Optimisation Example

	β_0	β_1	β_2		Fitness (MSE)
Child 1	-3.3745	5.2219	12.6500		0.2918 (say)

Next, apply a mutation operator to the child 'solution' and determine its fitness

Mutation allows for the discovery of information not contained in either parent

A Simple Optimisation Example

	$ß_0$	$ß_1$	$ß_2$		Fitness (MSE)
Child 1	-3.3745	5.2219	12.6500		0.2918 (say)
…	…	…	…		…
Child n	…	…	…		…

- Process is repeated until 'n' children are created
- These 'n' children form the next 'generation' of the population, and the algorithm continues
- Iteratively over time, the quality of members of the population improve and converge on the optimal values of $ß_0$, $ß_1$, $ß_2$

Crossover

- Crossover
  - Aims to use information from better parents ….
  - Could implement 'intermediate crossover' (a simple average)

Parent 1	-3.1245	5.6219	11.3411
Parent 2	-3.6245	4.8219	13.3411
Child	-3.3745	5.2219	12.3411

  - More generally, could use $P_1+\alpha(P_2-P_1)$, where P_1 and P_2 are the real-values in that locus of each parent and α is a scaling factor (perhaps randomly drawn from [-2, +2])
    - Defines a hypercube based on the current location of the parents

Mutation

- Mutation
  - Allows the uncovering of new information that was not present in either parent
  - Could add a random draw from $N(0, \alpha_i)$ to each element of each child solution
    - Hence, most mutations are small with occasional larger mutation steps
    - Value of α_i is user-defined (scaled as appropriate)

Model Calibration

- Often in finance, we may have a 'theory' / model which explains returns / prices etc.
- Model may be complex / non-linear
- In order to apply the model, its parameters need to be determined or 'calibrated'
  - Select model parameter values so that 'model output' matches actual market output
  - Calibrated model can then be used to (for example) price financial instruments

Model Calibration

Black Scholes Option Pricing Model

$$C_{BS}(S_t, K, r, q, \tau; \sigma) = S_t e^{-q\tau} N(d_1) - K e^{-r\tau} N(d_1)$$

$$d_1 = \frac{-\ln m + \left(r - q + \frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}} \qquad d_2 = d_1 - \sigma\sqrt{\tau}$$

Calibration (implied volatility)

$$\sigma_{BS}(S_t, K) > 0$$
$$C_{BS}(S_t, K, r, q, \tau; \sigma_{BS}(S_t, K)) = C_M(S_t, K)$$

Dang, J., Brabazon, A., O'Neill, M. and Edelman, D. (2008). Estimation of an EGARCH Volatility Option Pricing Model using a Bacteria Foraging Optimisation Algorithm, in Natural Computation in Computational Finance, Brabazon, A. and O'Neill, M. (eds), pp. 109-131, Berlin: Springer.

Fan, K., O'Sullivan, C., Brabazon, A., O'Neill, M. and McGarraghy, S. (2008). Calibration of the VGSSD Option Pricing Model using a Quantum Inspired Evolutionary Algorithm, in Quantum-Inspired Evolutionary Computation, Nedjah, N., Coelho, L. and Mourelle, L. (eds), pp. 133-153, Berlin: Springer.

Fan, K, O'Sullivan, C., Brabazon, A. and O'Neill, M. (2008). Non-linear Principal Component Analysis of the Implied Volatility Smile using a Quantum-inspred Evolutionary Algorithm, in Natural Computation in Computational Finance, Brabazon, A. and O'Neill, M. (eds), pp. 89-108, Berlin: Springer.

Parameterising a Simple Rule

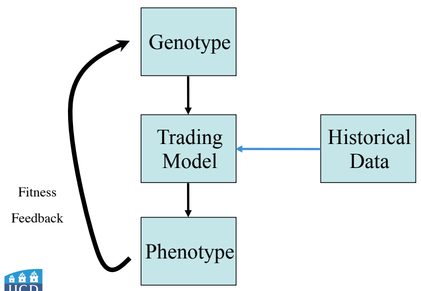
- How might you represent a simple technical trading rule of the following form as a string?

IF x day MA of price ≥ y day MA of price
THEN Go Long *ELSE* Go Short

5	10	1

(*5 day MA* ≥ 10 day MA THEN Go Long)

Back-testing the Quality of a Genotype …



UCD Natural Computing Research & Applications Group (NCRA)

25

A Slightly More Complex Example …

IF [Indicator$_i$(t) (<,>) value$_j$] *THEN* (Buy, Sell, Do nothing)

Indicator$_i$	t	<,>	value$_j$	Buy, Sell, Do nothing

- The above are simple illustrative examples, much more complex, compound, trading rules, which would defy any attempt at discovery via enumerative methods, could also be generated using AND, NOT, OR etc. operators

UCD Natural Computing Research & Applications Group (NCRA)

26

Uncovering an Stock Selection Rule

High sales growth relative to industry average?	High debt level relative to industry average?	High level of cash flow from operations relative to industry average?	High level of liquidity relative to industry average?	High profit level relative to industry average?

In a simple case, we may be trying to uncover a good subset from an array of plausible filter rules (possible rules depend on your investment style)

Each indicator could be coded as a 0 (no) or 1 (yes), with an evolutionary process being applied to uncover the best subset of filter rule components

Filter 1	…			Filter n	R/σ (say)
0	…	0	1	1	XX

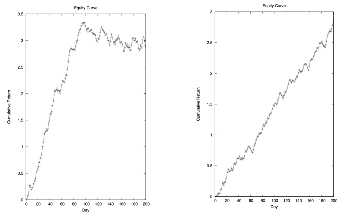
Of course, you could also apply an evolutionary process to breed the individual elements of the filter rules and their thresholds … but this is better done using an evolutionary model induction methodology

UCD Natural Computing Research & Applications Group (NCRA)

27

Generalisation / Defining an Appropriate Fitness Function

A statistical measure of 'goodness of fit' is not the same thing as 'profitability'!



Bradley, R., Brabazon, A. and O'Neill, M. (2010). Objective Function Design in a Grammatical Evolutionary Trading System, Proceedings of the 2010 IEEE Congress On Evolutionary Computation, IEEE Press.

Agapitos, A., O'Neill, M. and Brabazon, A. (2010). Evolutionary Learning of Technical Trading Rules without Data-mining Bias, Proceedings of 11th International Conference on Parallel Problem Solving from Nature (PPSN 2010), Lecture Notes in Computer Science, Springer-Verlag, Berlin.

UCD Natural Computing Research & Applications Group (NCRA)

28

Model Induction

- Can we 'recover' a suitable model (+ associated parameters) from the data?
  - ANNs (universal approximators … but readability?)
  - GP / GE …
  - Likely to be useful when we have data but weak theory (perhaps some idea of the likely relevant variables but little idea how they might link together)
- Applications include:
  - financial time-series forecasting, credit risk modelling, pricing model discovery, forecasting takeover targets, prediction of earnings, IPO underpricing, trading system etc. etc.

UCD Natural Computing Research & Applications Group (NCRA)

29

Genetic Programming

- An evolutionary model-induction methodology
- Idea dates from at least the 1950s, popularised by John Koza

  in his 1992 book '*Genetic Programming: on the programming of computers by means of natural selection'*
- GP adopts an evolutionary metaphor
  - Generate a population of trial solutions, assess worth of each, select, crossover, mutate, replace

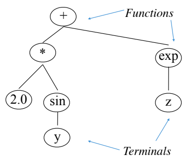
UCD Natural Computing Research & Applications Group (NCRA)

30

Genetic Programming

- Individual is or represents a program

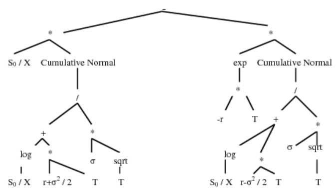
```
#include<stdio.h>
#include<stdlib.h>
…
int main(int argc, char* argv){
    float x=0.0, y=0.0, z=0.0;
    x=atof(argv[1]); y=atof(argv[1]); z=atof(argv[1]);
    z1 = 2.0*sin(y) + exp(z);
    printf("The answer is: %f\n",z1);
    return (0);
}
```



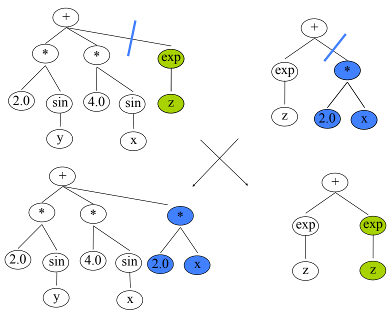
Of course, a 'program' (or the equivalent tree representation) can be viewed as a list of rules … and many financial problems can be viewed as a search for a 'good' list of rules …. lending decisions, investment decisions, ….

Genetic Programming

Stylised Tree representation of the BS option pricing model



Diversity Generation



Genetic Programming

- Typically, in financial applications of GP, the goal is to recover / discover the data-generating model
  - What model can we reverse engineer from the data?
  - Utility in building forecasting models … but also in theory development …
- As each 'model' is evolved, it's quality / fitness can be assessed by determining how well it explains the observed (training) data

GP Pseudo-Code

Define terminals, primitive functions and fitness function
Set parameters for GP run (population size, probabilities for mutation, crossover etc., selection / replacement strategy etc.)
Initialise start population of solutions (grow, full, ramped-half and half)
Calculate fitness of each solution (run each program!)
WHILE (Termination condition False)
    Select parents
    Create offspring
        using mutation, crossover, cloning, architecture-altering...
    Update population
    Calculate fitness of each solution
ENDWHILE

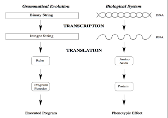
Open Issues in GP

O'Neill, Vanneschi, Gustafson, Banzhaf. (2010). *Open Issues in Genetic Programming*. Genetic Programming & Evolvable Machines, 11(3).
GP&EM 10th Anniversary Issue

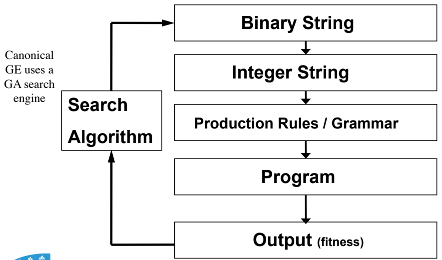
- Identifying appropriate Representations
- Fitness Landscapes & Problem Difficulty
- Static vs. Dynamic Problems
- The Influence of Biology
- Open-ended Evolution
- Generalization
- Benchmarks
- Modularity
- Complexity of GP
- Misc....
  - Halting, AI Ratio, Bio, Constants, Theory, Distributed Models, Usability...

Grammatical Evolution

- Grammar-based GP … model specification / choice of explanatory variables not assumed ex-ante
- Genotype → phenotype mapping
  - Each program is generated from a variable length linear binary (or integer) string
  - Key item is that the evolutionary process is effectively applied to the 'production rules'
    - i.e. the developmental rules governing the production of the phenotype, and not directly to the phenotype itself

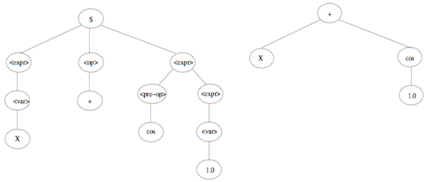


Grammatical Evolution



Genetic Operators

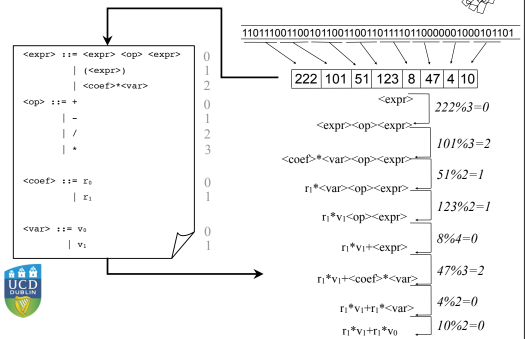
- (Variable-length) Binary/Integer String
- Bit/Codon Mutation
- 1pt Xover
- Duplication
- Tree-based operators



BNF Grammar

```
T = {sin, cos, tan, log, +, -, /, *, x, (, )}
S = <expr>
N = {<expr>, <op>, <pre-op>, <var>}
P =
        <expr> ::= (<op> <expr> <expr>)
                 | (<pre-op> <expr>)
                 | <var>

        <op> ::= + | - | / | *

        <pre-op> ::= sin | cos | tan | log

        <var> ::= x
```


Example Mapping

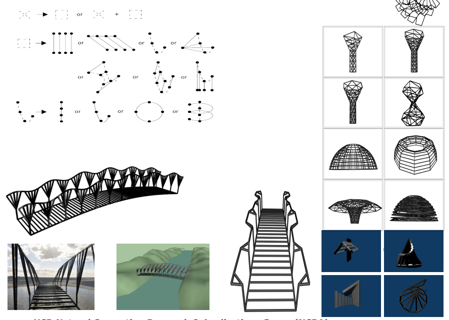


Sample Grammars

```
<lc> ::= output = <expr> ;

<expr> ::= ( <expr> ) + ( <expr> )
         | <coeff> * (<var>/<var>)

<var> ::= var1[index] | .... | var12[index]

<coeff> ::= (<coeff>) <op> ( <coeff> )
          | <float>

<op> ::= + | -

<float> ::= 20 | -20 | 10 | -10 | 5 | -5 | 4 | -4
          | 3 | -3 | 2 | -2 | 1 | -1| .1 | -.1
```

Output = coeff * (var 1 / var 3) + ···

Shape Grammars

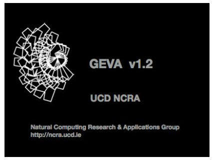


Want to learn more?

Released 1st July.....

http://ncra.ucd.ie/geva/

O'Neill, Hemberg, Gilligan, Bartley, McDermott, Brabazon. (2008). *GEVA: Grammatical Evolution in Java*. SIGEVOlution, 3(2):17-23.



Dempsey, O'Neill, Brabazon. (2009). ***Foundations of GE in Dynamic Environments***. Springer.

McKay, Nguyen, Whigham, Shan, O'Neill. (2010). *Grammar-based Genetic Programming - A Survey*. Genetic Programming & Evolvable Machines 11(3).

Brabazon, O'Neill. (2006). *Biologically Inspired Algorithms for Financial Modelling*. Springer.

O'Neill, Ryan. (2003). *Grammatical Evolution*. Kluwer Academic Press.

Financial Applications …

- Recent Financial Applications of GE include:
  - Asset selection
  - Options price modelling
  - Credit risk assessment
  - Algorithmic trading
  - Money-laundering detection …..

Credit Risk Modelling
Brabazon, A. and O'Neill, M. (2006). Credit Classification Using Grammatical Evolution. Informatica, 30(3):325-335.
Brabazon, A. and O'Neill, M. (2008). Bond Rating with piGrammatical Evolution, in Studies in Computational Intelligence: Knowledge-Driven Computing, Knowledge Engineering and Intelligent Computation, Cotta, C., Reich, S., Ligęza, A. and Schafer, R. (eds), pp. 17-30, Berlin: Springer Verlag.

Corporate Failure Prediction
Brabazon, A. and Keenan, P. (2004). A hybrid genetic model for the prediction of corporate failure, Computational Management Science, 1 (3-4):293-310.
Brabazon, A. and O'Neill. M. (2004). Diagnosing Corporate Stability using Grammatical Evolution, International Journal of Applied Mathematics and Computer Science, 14(3):363-374.

Trading System Generation
Brabazon, A. and O'Neill, M. (2004). Evolving technical trading rules for foreign-exchange markets, Computational Management Science, 1 (3-4):311-327.
Brabazon, A., Meagher, K., Carty, E. , O'Neill, M. and Keenan, P. (2005). Grammar-mediated time-series prediction, Journal of Intelligent Systems, 14(2-3):123-143.

Bond Rating Example

- Companies can be financed by share capital or debt capital
- Debt capital → Bank debt; → Bond issue
- If bonds are to be publicly traded, they require (in US) a credit rating from a 'recognised' rating agency (S&P, Moody's, Fitches') …
  S&P cover 99.2% of traded debt in the US
- Credit rating represents an agency's opinion of the creditworthiness of a borrower

Background

Investment Grade

Initial Rating	Default (%)
AAA	0.52
AA	1.31
A	2.32
BBB	6.64
BB	19.52
B	35.76
CCC	54.38

Source: S&P 1987-2002

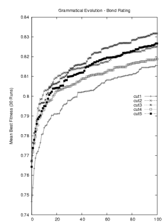
Results

Grammar

```
<lc> ::= if( <expr> <relop> <expr> ) class="Junk";
      else class="Investment Grade";
<expr> ::= ( <expr> ) + ( <expr> ) | <coeff> * <var>
<var> ::= var1[index] |… ] | var8[index]
<coeff> ::=  ( <coeff> ) <op> ( <coeff> ) | <float>
<op> ::= + | - | *
<float> ::= 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | -1 | .1
<relop> ::= <=
```

Data
600 US firms (420 training:180 out of sample)

S&P bond ratings + data drawn from their financial statements



Best fitness values, over 30 runs and five recuts, on the training data (popsize=500)

Coefficients concord with perceived financial wisdom …

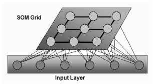
Low/negative retained earnings; low / negative total assets; high levels of debt, are symptomatic of firms with a junk rating

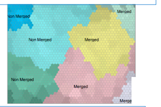
Model Induction

Other Applications / Other Methods

Hickey, R, Little, E. and Brabazon, A. (2006). Identifying Merger and Takeover Targets Using a Self-Organising Map, Proceedings of the 2006 International Conference on Artificial Intelligence (ICAI '06), edited by Arabnia, H. et al., Las Vegas 26-29 June 2006, (ISBN), Vol. 1, pp. 408-413: CSEA Press.

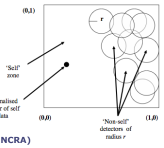
Le-Khac,, N. A., Markos, S., O'Neill, M., Brabazon, A. and Kechadi, M. T. (2009). An investigation into Data Mining approaches for Anti-money Laundering, IEEE International Conference on Knowledge Discovery, (ICKD'09), Manila, Philippine, 6-8 June 2009, IEEE Press.





Cahill, J., Keenan, P., Walsh, D. and Brabazon, A. (2010). Identifying Online Credit Card Fraud using Artificial Immune Systems, Proceedings of the 2010 IEEE Congress On Evolutionary Computation, IEEE Press.

Brabazon, A., Delahunty, A., O'Callaghan, D., Keenan, P. and O'Neill, M. (2007). Financial Classification using an Artificial Immune System, in Intelligent Information Technologies: Concepts, Methodologies, Tools, and Applications, Sugumaran, V. (ed), pp. 1525-1540, Hershey, PA, USA: Information Science Reference.



Model Induction

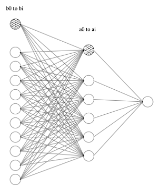
Prediction of Going Concern Audit Qualification

Data:
•Liquidity, Debt, Profitability, Activity / Efficiency, Cash Generation

•Rate of change of above

•Loss in current year, Loss in past three years, Big-six auditor

MLP	Train	Out of sample
T-1	72.79%	70.61%
T-2	68.05%	64.78%
T-3	65.38%	57.57%



Thompson, D., Thompson, S. and Brabazon, A. (2007). Predicting Going Concern Audit Qualification using Neural Networks, Proceedings of the 2007 International Conference on Artificial Intelligence (ICAI '07), edited by Arabnia, H. et al., Las Vegas 25-28 June 2007, Vol. 1, pp. 199-204, CSEA Press.

The Trading Environment

BATS MSFT Go

MICROSOFT CORP COM

Orders Accepted 136,272   Total Volume 6,305,106

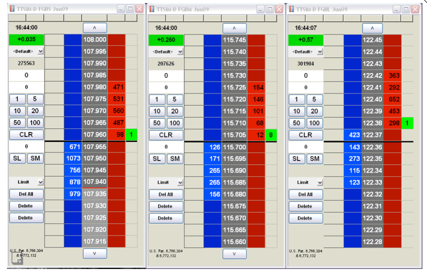
	TOP OF BOOK		LAST 10 TRADES		
	SHARES	PRICE	TIME	PRICE	SHARES
ASKS	14,500	23.8500	14:19:43	23.8000	100
ASKS	14,600	23.8400	14:19:43	23.8000	100
ASKS	6,500	23.8300	14:19:41	23.8050	200
ASKS	21,300	23.8200	14:19:40	23.8000	200
ASKS	6,800	**23.8100**	14:19:40	23.8000	100
BIDS	7,900	**23.8000**	14:19:40	23.8000	1,000
BIDS	4,604	23.7900	14:19:39	23.8000	100
BIDS	7,800	23.7800	14:19:34	23.8050	200
BIDS	11,100	23.7700	14:19:20	23.8000	1,400
BIDS	11,600	23.7600	14:19:09	23.8100	100

Last updated 14:19:48

Snapshot of BATS order book for Microsoft 8 September 2010

Source: http:batstrading.com/

The Trading Environment

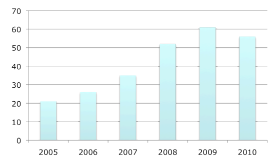


The Trading Environment

- Growth of Algorithmic Trading
  - Can encompass any aspect of the trading process
    - Trade execution only, or …
    - Entire process of opportunity identification and trading e.g. arbitrage trading
  - Fragmentation of Liquidity
  - Algorithmic Gaming ….

Algorithmic Trading

- High-frequency trading (HFT)
  - Algorithm may place 000s of orders per second (with up to 90% being cancelled)



Percentage of US turnover by volume accounted for by HFT (Source: Tabb Group, 2010)

Algorithmic Trading

- High-frequency trading (HFT)
  - Concurrent need for low latency both in terms of algorithmic decision-making (opportunity identification and execution), data link speed (co-location with exchange matching engine is now common … NYSE moved to new data centre in NJ at end-AUG), and speed of processing of exchange platform to incoming orders

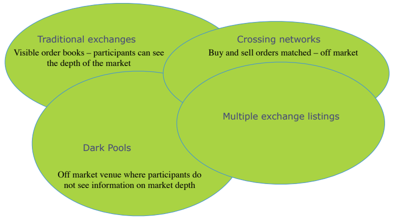
Exchange	Time to execute a trade (milliseconds)
NASDAQ	0.177
London / Sydney	3
Tokyo / Bombay	5
Hong Kong	9
Brazil	15
Singapore	16

Source: Mondo Visione (2010)

  - At these speeds, even internal machine processes can become bottlenecks e.g. system processes which clean house and reassign memory can slow down market response … (hence, HF Traders may look to 'control' when these processes occur ….!)

Algorithmic Trading

Fragmentation of liquidity…. e.g. only 55% of trading in FTSE 100 stocks now occurs in London …



Algorithmic Trading

- Algorithmic gaming
  - 'Phantom liquidity'
  - 'Fishing'
    - Display an order on one side of the book in order to coax liquidity onto the other side … then cancel the initial order and hit the newly

Algorithmic Trading

- Algorithmic gaming
  - 'Quote stuffing'
    - Send heavy order traffic to exchange which are almost immediately cancelled in order to slow down data feeds from the exchange
      - e.g. a long CQS quote is 102 bytes. Suppose an algorithm generates 5,000 orders / cancellations per second for 1 stock … this generates (5,000 x 102) 510,000 bytes per second message traffic from the exchange
      - As a T1 line has a capacity of 154,000 bytes/sec, hence the above quote stuffing would consume 3 T1 lines, of course, you could have DS3 bandwidth … !!
    - This opens up microsecond arbitrage opportunities between different venues as the 'quote stuffer can differentiate the true liquidity from the false whereas competing algorithms become overloaded. It also causes competitors' algorithms to make decisions based on faulty data

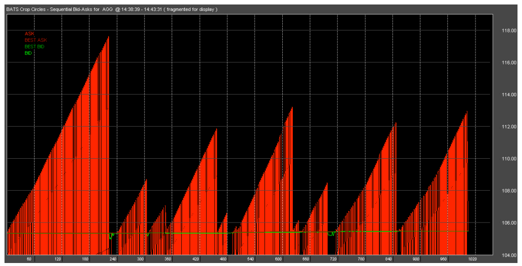
Algorithmic Trading

Flash Crash 6 May, 2010 - DJIA drops 1000 points in 20 minutes mid session

Source: http://en.wikipedia.org/wiki/File:Chart_dow_dip2.top.gif

Estimated that this cost retail investors >$2bn as stop-loss orders were hit

Algorithmic Trading



The Bandsaw ….. Unusual trading activity on BATs (6 May 2010), (AGG - EFT based on US investment grade bonds)

Source: http://www.nanex.net/20100506/FlashCrashAnalysis_Part4-1.html

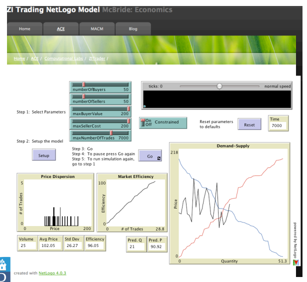
Agent-based Modelling of Markets (ASM)

- Actual market history is a single-sample path
  - Observation does not allow 'what-if' analyses
- ABM adopts a 'bottom up' approach to 'simulate' financial markets (markets 'in silico')
  - Autonomous agents whose interactions produce complex, emergent, system dynamics
    - Can embed heterogeneous agents with differing risk attitudes, with differing expectations to future outcomes, and with learning/adaptive capabilities
    - How do agents learn? (application of NC … GP, NN … etc.)
    - The 'Red Queen'

Agent-based Modelling of Markets (ASM)

- Plausibility testing of ABM / Model validation (?)
  - Can we replicate observed real world characteristics of markets?
  - What are the critical assumptions?
- Applications of ABM
  - Attempt to explain market behaviour (theory building)
  - Provide insights for policy makers and regulators
  - Provide test-bed for 'simulation' of trading strategies / execution strategies

Agent-based Modelling of Markets (ASM)



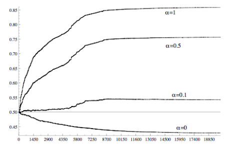
Source: Mark E. McBride
Economics , Miami University

Model based on Gode and Sunder (1993)

http://mcbridme.sba.muohio.edu/ace/labs/zitrade/zitradenetlogo.html

ABM for Understanding

- Market fraction asset pricing model in order to investigate the market dominance, the profitability, and the survival rates of both fundamental and trend-following investors across varying time scales
- Simulation results results indicate that in contrast to the prediction of traditional financial theory, trend-followers can survive in the market in the long run and in the short run they can outperform fundamentalists

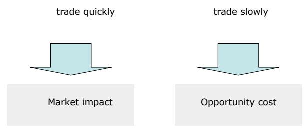


Time series of absolute wealth accumulation of the fundamentalists for varying levels of alpha (confidence of fundamentalists in their estimates of fundamental value). Note, trend followers – technical traders - persist (survive) at all levels of alpha.

He, X-Z., Hamill, P. and Li, Y. (2008) Can Trend Followers Survive in the Long-Run?: Insights from Agent-Based Modeling, , in Natural Computation in Computational Finance (Volume I), Brabazon, A., O'Neill, M. (eds), pp. 253-269, Berlin: Springer.

ASM as a Test Bed

- How can investors buy/sell large orders efficiently?
- The *traders dilemma ….*



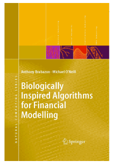
- We need to balance *Market Impact* (embeds cost of 'liquidity demand' and 'information leakage') and *Opportunity cost*

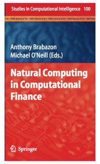
Trade Execution Strategy

- Trade execution requires the 'design' of an appropriate 'trading strategy'
  - Order scheduling (number of sub-orders, their size, when submitted to market)
  - Order aggressiveness (what 'style' to adopt?)
  - Split of order across markets
- A practical problem then is how do we assess the utility of a strategy without actually implementing it?
- Scope to 'test' in an ASM

Cui, W., Brabazon, A. and O'Neill, M. (2010). Evolutionary Computation in Trade Execution, in Natural Computation in Computational Finance (Volume III), Brabazon, A., O'Neill, M. and Maringer, D. (eds), pp. 45-62, Berlin: Springer.

References





Brabazon, A. and O'Neill, M. (2006). Biologically Inspired Algorithms for Financial Modelling, Berlin: Springer.

Brabazon, A. and O'Neill, M. (eds) (2008) (2009) (2010). Natural Computation in Computational Finance, Berlin: Springer-Verlag.

(Vols 1-3)

UCD Natural Computing Research & Applications Group (NCRA)

evo* 2011

CALL FOR PAPERS

EvoFIN 2011
5th European event on Evolutionary and Natural Computation in Finance and Economics
27-29 April 2011
Torino, Italy

EvoFIN 2011 focuses primarily on the use of EC and related Natural Computing techniques in Computational Finance and Economics.

Applications of interest include (but are not limited to) forecasting financial time series, portfolio selection and management, estimating econometric parameters, pricing options, and developing risk management systems. Other applications of interest include artificial stock markets.

Methodologies of interest include EC methods such as genetic programming, genetic algorithms and evolutionary strategies, as well as other related natural computing methodologies such as particle swarm, foraging algorithms, artificial immune systems, hybrid systems and agent-based systems.

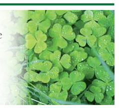
Conference proceedings are published as part of Springer's LNCS series

Submission deadline: 22nd November

http://evostar.dei.uc.pt/

Call for Papers

2011 Genetic and Evolutionary Computation Conference
Tuesday – Saturday July 12–16, 2011 Dublin, Ireland
Largest Conference in the Field of Genetic and Evolutionary Computation
A recombination of the
20th International Conference on Genetic Algorithms (ICGA) and the
16th Annual Genetic Programming Conference (GP)
www.sigevo.org/gecco-2011



One Conference – Many Mini-Conferences

Program Tracks

Three days of presentations of the latest high-quality results in more than 15 separate and independent program tracks specializing in various aspects of genetic and evolutionary computation.

Free Tutorials and Workshops!

Two days of free tutorials and workshops (included with conference registration) presented by some of the world's foremost experts in topics of interest to genetic and evolutionary computation researchers and practitioners.

Important Dates

* **Submission deadline: January 26, 2011**
* Notification of paper acceptance: March 23, 2011
* Camera-ready submission: April 18, 2011
* Workshop and tutorial proposals submission: November 8, 2010
* Notification of workshop and tutorial acceptance: November 29, 2010
* GECCO-2011 Conference: July 12-16, 2011

How to Submit a Paper

Meet the submission deadline (January 26, 2011) and submit substantially new work. GECCO allows submissions of material that is substantially similar to a paper being submitted contemporaneously for review in another conference. However, if the submitted paper is accepted by GECCO, the authors agree that substantially the same material will not be published by another conference in the evolutionary computation field. Material may be later revised and submitted to a journal, if permitted by the journal.

Visit www.sigevo.org/gecco-2011

For detailed instructions, including categories, keywords, and formatting requirements, visit http://www.sigevo.org/gecco-2011. Be sure to check the web page for changes that may appear as the paper submission deadline approaches.

Follow us at www.twitter.com/GECCO2011

News and updates are also posted on our Twitter page. Visit http://www.twitter.com/GECCO2011 and start following us!

Accept author agreement

By submitting a paper, the author(s) agree that, if their paper is accepted, they will:

* Submit a final, revised, camera-ready version to the publisher by Monday, April 18, 2011
* Register at least one author to attend the conference by