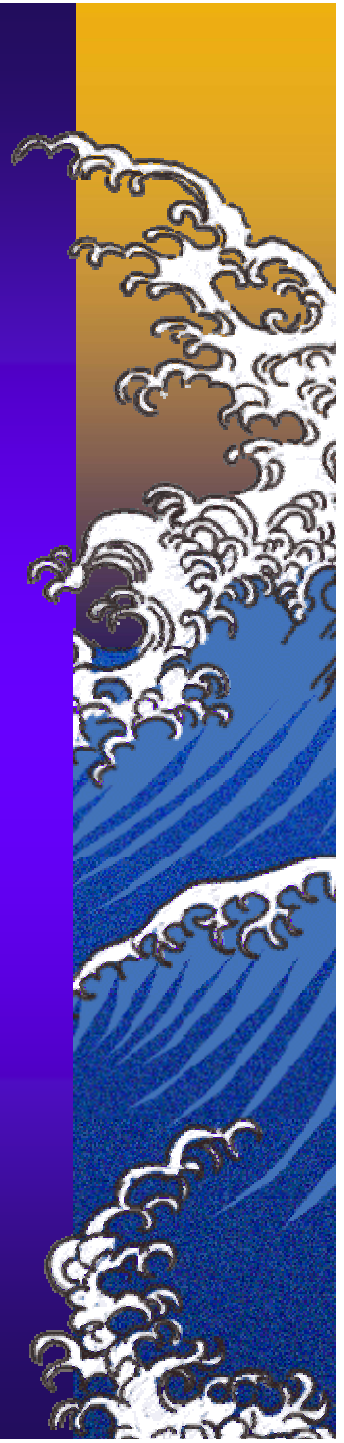
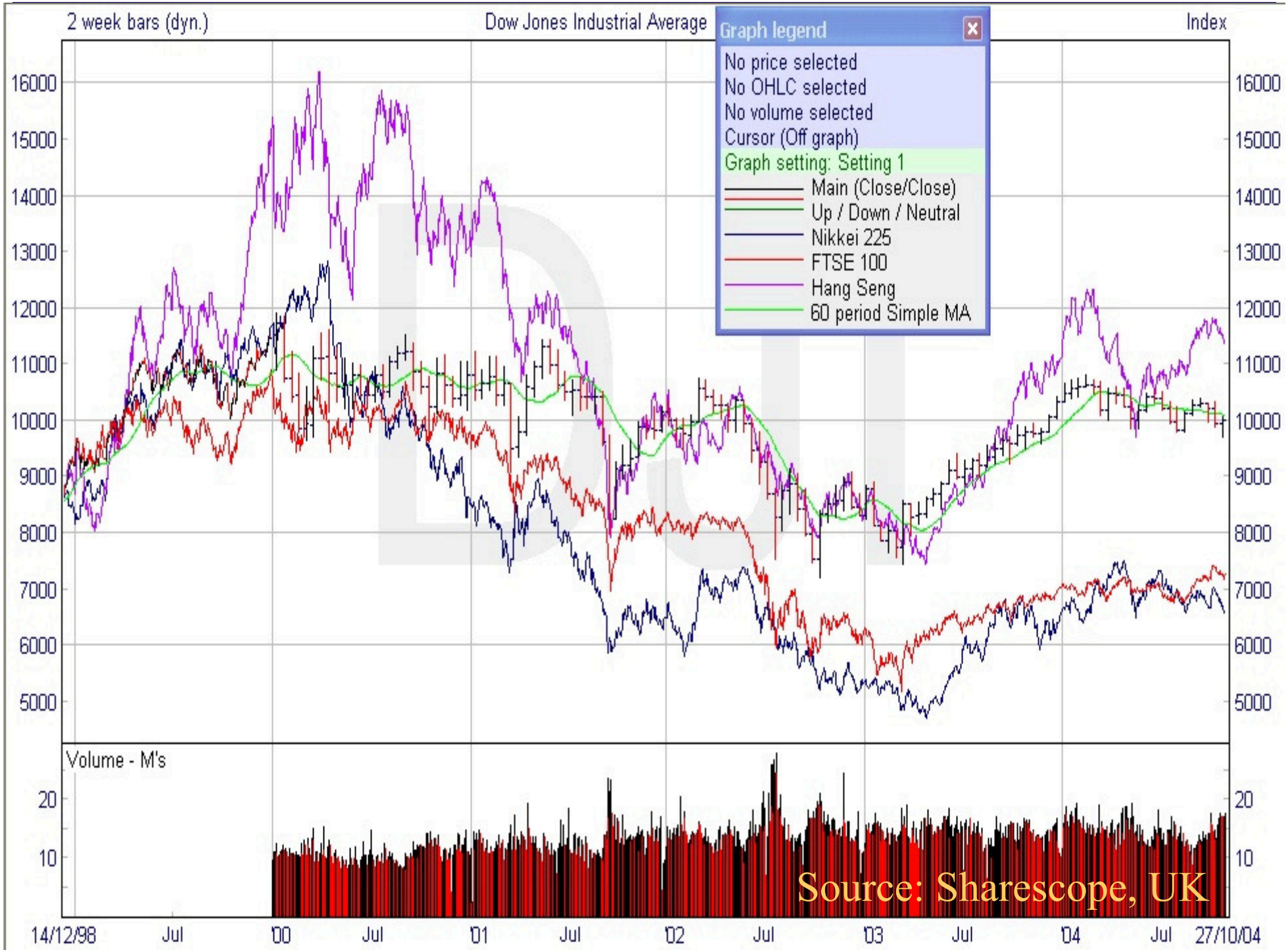
EDDIE:
A Genetic Programming Tool for Financial Forecasting

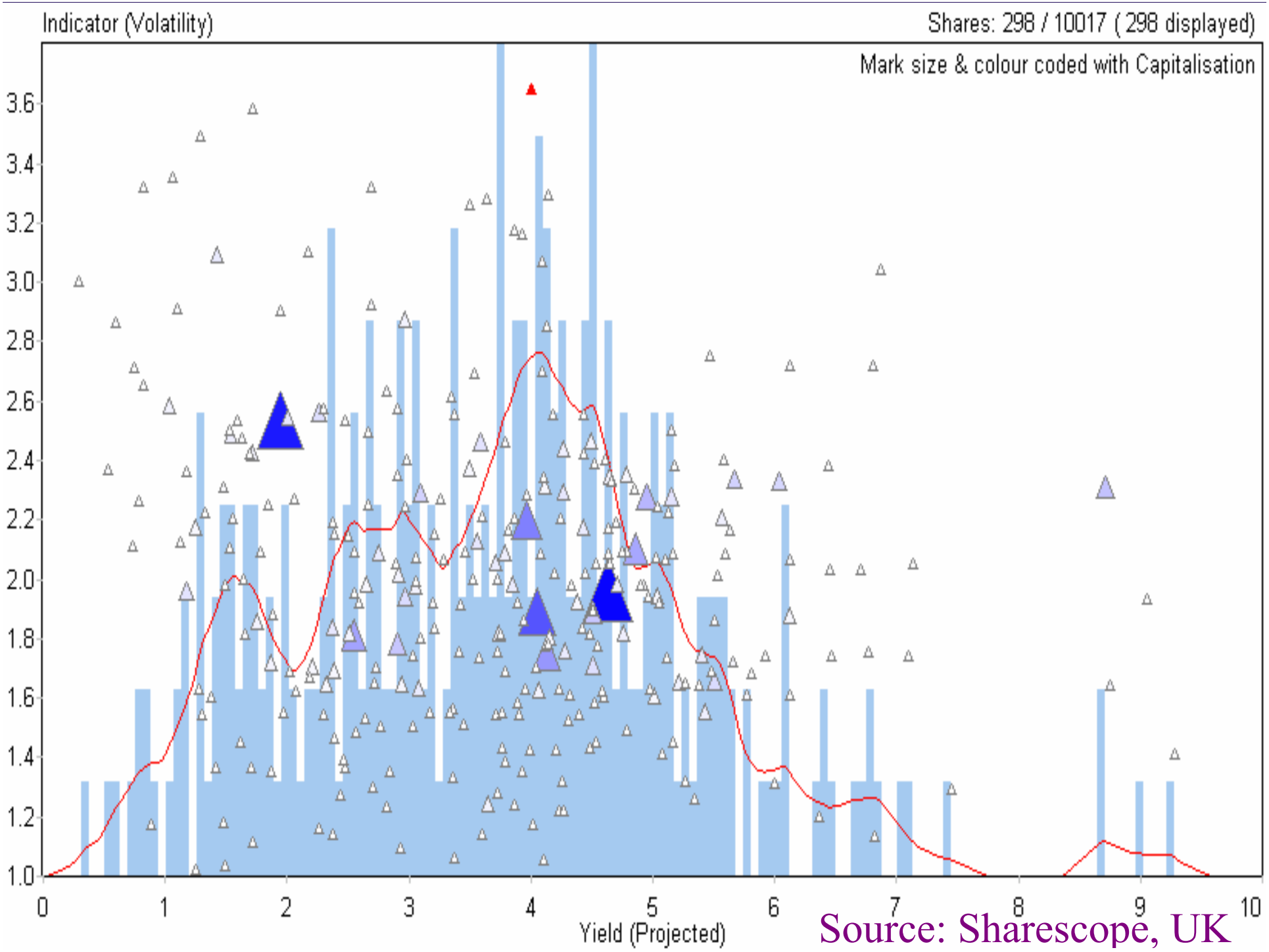
Edward Tsang

University of Essex

EDDIE = **E**volutionary **D**ynamic **D**ata **I**nvestment **E**valuator




2 week bars (dyn.)
Dow Jones Industrial Average
Index
Graph legend
No price selected
No OHLC selected
No volume selected
Cursor (Off graph)
Graph setting: Setting 1
Main (Close/Close)
Up / Down / Neutral
Nikkei 225
FTSE 100
Hang Seng
60 period Simple MA
16000
15000
14000
13000
12000
11000
10000
9000
8000
7000
6000
5000
Volume - M's
20
10
14/12/98
Jul
'00
Jul
'01
Jul
'02
Jul
'03
Jul
'04
Jul
27/10/04
Source: Sharescope, UK



Source: Sharescope, UK

EDDIE Research Agenda

- Suppose your expert tells you that
  - Price-earning ratio
  - 12 or 50-days moving average
  - Interest rate
  - ….

  are relevant to future price of FTSE-100
- How would you actually use them to forecast?
- Goal: add value to your expert knowledge

Efficient Market Hypothesis (EMH)

- Financial assets (e.g. shares) pricing:
  - All available information is discounted
- If EMH holds, forecasting is impossible
  - Random walk theory
- Assumptions:
  - Efficient markets
  - Perfect information flow
  - Rational traders

Does the EMH Hold?

- It holds for the long term
- "*Fat Tail*" observation:
  - big changes today often followed by big changes (either + or –) tomorrow
- How fast can one adjust asset prices given a new piece of information?
  - Faster machines certainly help
  - So should faster algorithms

EDDIE / FGP Overview

EDDIE = **E**volutionary **D**ynamic **D**ata **I**nvestment **E**valuator
FGP = **F**inancial **G**enetic **P**rogramming

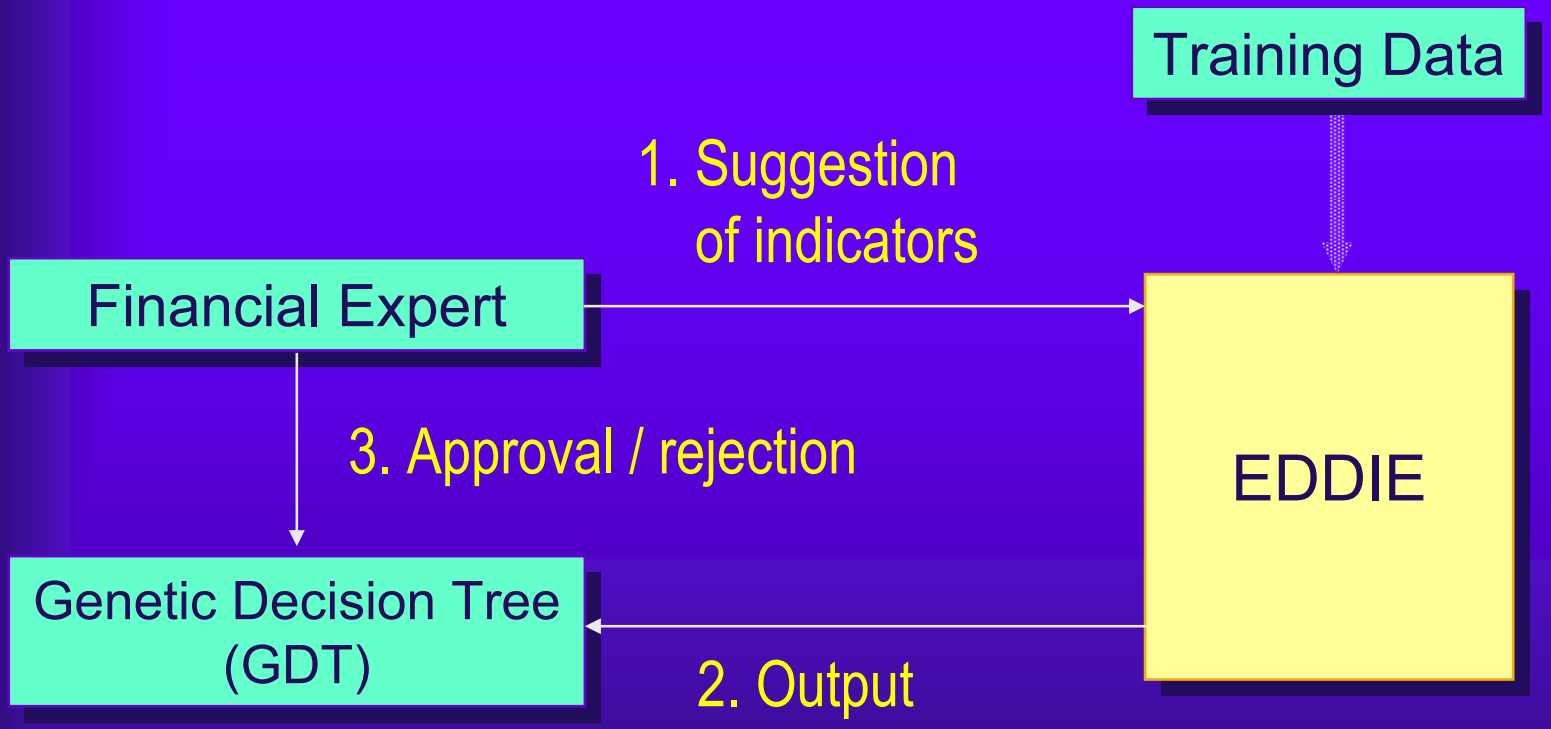
James Butler

- EDDIE / FGP learns from past history
  - Using Genetic Programming
- It generates decision trees
  - Which allows it to *explain* its recommendations
- Used learned rules to answer questions such as:
  - Will prices rise by 4% within the next 21 days?
- It works with *domain experts* on
  - what *features* are relevant?
  - are the rules generated *reasonable*?

Working with Experts

- EDDIE is not designed to *replace* experts
  - It is designed to work *with* experts
- GP is only a tool
  - it needs expert input to be effective
- Experts channel knowledge into EDDIE:
  - by suggesting what factors are relevant
  - by evaluation of the rules generated
- EDDIE adds value expert input

Expert Knowledge in EDDIE

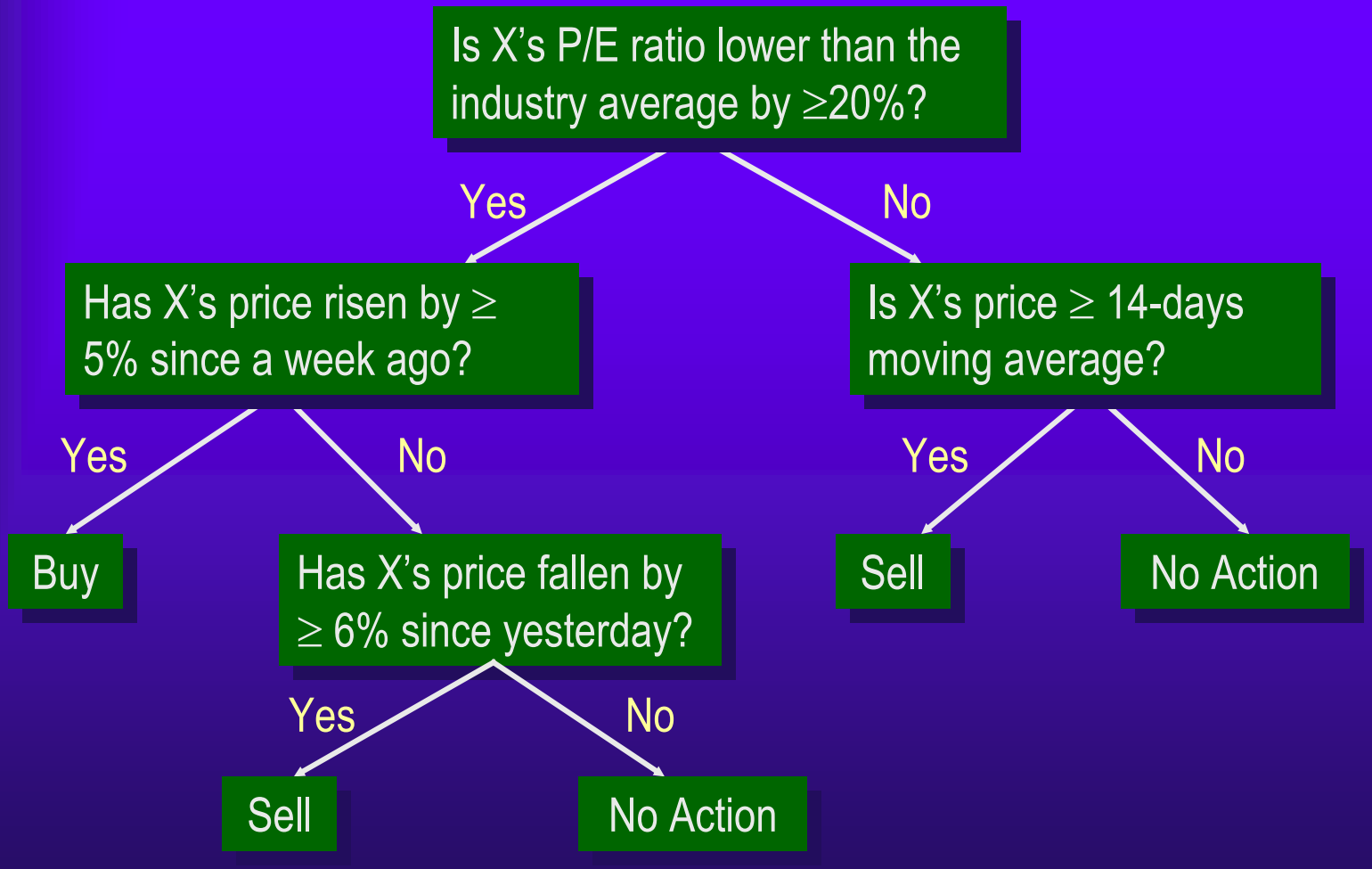


- Effective channeling of expert knowledge into EDDIE is the key to success

Technical Details

Inside EDDIE / FGP

An Example Decision Tree



Syntax of GDTs in EDDIE-2

<Tree> ::= “If-then-else” <Condition> <Tree> <Tree> | Decision

<Condition> ::= <Condition> "And" <Condition> |

<Condition> "Or" <Condition> |

"Not" <Condition> |

Variable <RelationOperation> Threshold

<RelationOperation> ::= ">" | "<" | "="

Variable is an indicator / feature

Decision is an integer, “Positive” or “Negative” implemented

Threshold is a real number

- Richer language ⇒ larger search space

A taste of user input

Given	Expert adds:	More input:		Define target:
Daily closing	50 days m.a.	Volat-ility	.....	↑4% in 21 days?
90	80	50		0
99	82	52		0
87	83	53		1
82	82	51		1
.....	.....	.....		.....

EDDIE adds value to user input

- User inputs *indicators*
  - e.g. moving average, volatility, predications
- EDDIE makes *selectors*
  - e.g. “50 days moving average > 89.76”
- EDDIE combines selectors into *trees*
  - by discovering interactions between selectors

➢ Finding thresholds (e.g. 89.76) and interactions by human experts is laborious

Research Methodology

- Concentrate on predicting:

  G = "will prices go up/down by *r*% within the next *n* days?"

- To evaluate EDDIE, choose *r* and *n* such that 50% of the days achieve G
  - Performance against random decisions
  - Also compared against ID3 / C4.5
- Measure prediction accuracy
  - Return on investment also used for reference

Testing of EDDIE

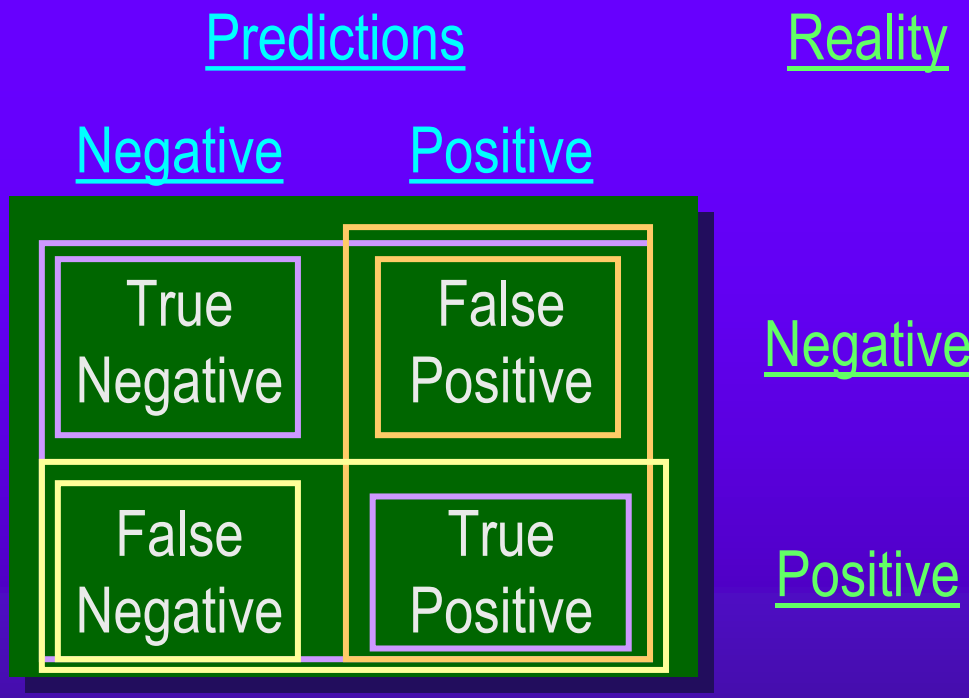
- S&P 500 Index, 1963 to 1974
- Dow Jones Industrial Average Index
- Combining expert predictions on Heng Seng Index
- Shares: IBM, HSBC, BAA, BHP, ANZ, 1991 to 2000
- UK handicap horse races 1993

EDDIE on S&P 500 daily close

- Trained: 2/4/63 to 2/7/70 (1800 days)
- Tested:6/7/70 to 25/1/74 (900 days)
- Target: "*rise of 4% within 63 days*"
- Input: textbook technical indicators
  - e.g. *n* days moving averages/ min/ max prices
- Result: 54% accuracy, 43% annual return

Reference: *Software, Practice & Experience*, 28(10) 1998

Performance Measures



- Rate of correctness = (TN + TP) ÷ Total
- Rate of failure = FP ÷ (FP + TP) = 1 – *precision*
- Rate of missing chances = FN ÷ (FN + TP) = 1 – *recall*

EDDIE on IBM 1991-1997

- 60% of recommendations correct
  - where opportunities occur in 45% of the days

IBM		Training Period (1991.10.30-1995.12.27)			
RC	61.4%	0	1	Opport.:	42.3%
RMC	89.1%	654	10	0	664
RF	15.9%	434	53	1	487
		1088	63	5%	1151
IBM		Over Test Period (1995.12.28-1997.03.05)			
RC	56.5%	0	1	Opport.:	45.0%
RMC	90.0%	104	6	0	110
RF	40.0%	81	9	1	90
AR	210.0%	185	15	8%	200

FGP2 on HSBC 1996-2000

- No recommendations made

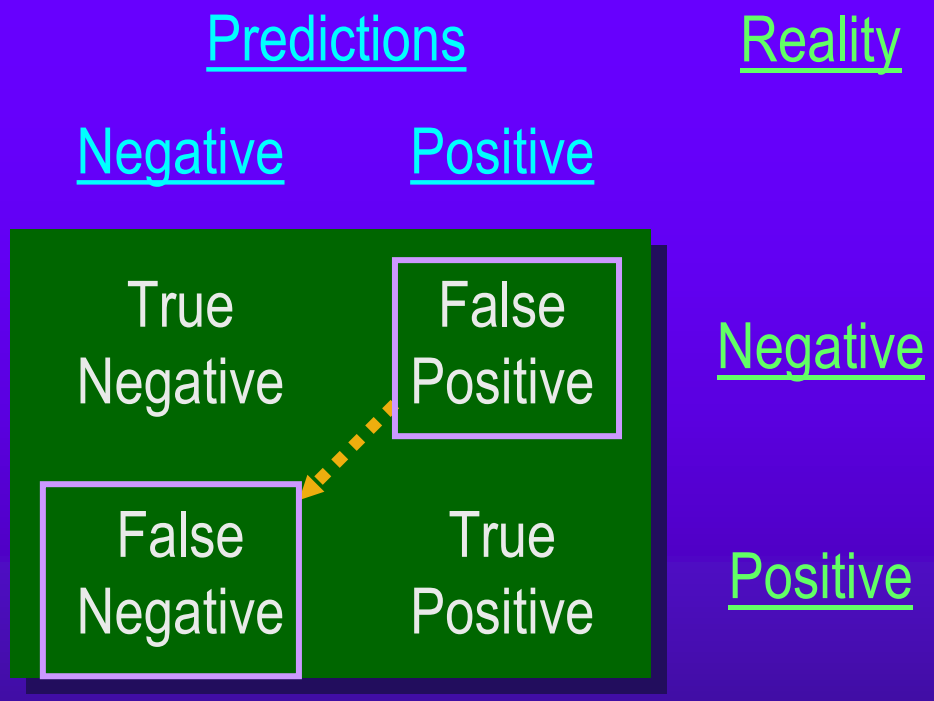
HSBA		Training	12/03/96	to	28/5/1999
RC	54.8%	0	1	Opport.:	52.1%
RMC	83.8%	389	13	0	402
RF	15.5%	366	71	1	437
		755	84	10%	839

HSBA		Testing	31/5/1999	to	03/03/00
RC	59.0%	0	1	Opport.:	41.0%
RMC	100.0%	118	0	0	118
RF	N.A.	82	0	1	82
		200	0	0%	200

Improving Precision

Reducing Rate of Failure

Reducing Rate of Failure (RF)



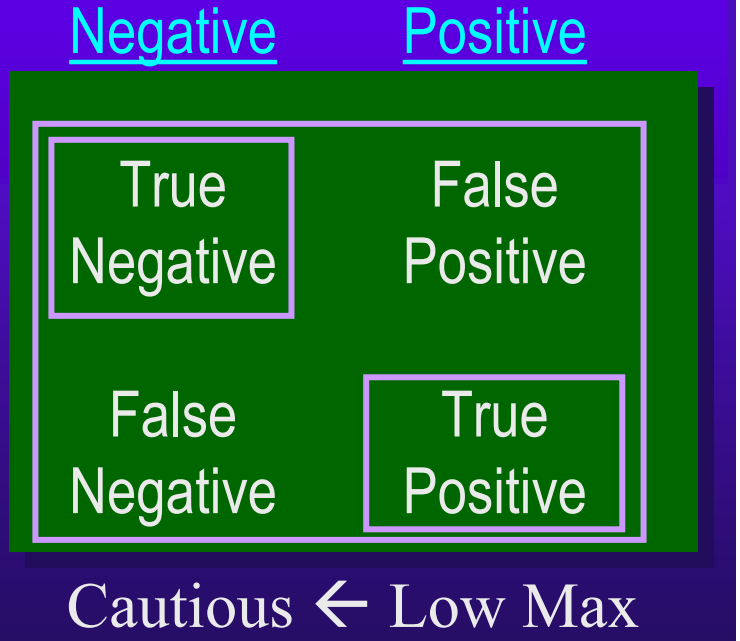
- RF = FP ÷ (FP + TP) = 1 – *precision*
- RMC = FN ÷ (FN + TP) = 1 – *recall*
- Reduce RF at the cost of RMC

FGP: Constrained Fitness

Jin Li
FGP

- Constraints can help guiding the search
- Fitness = $w_{rc} \times RC' - w_{rmc} \times RMC - w_{rf} \times RF$
- RC' = RC if P+ $\in$ [Min, Max]
  0 otherwise

- One can adjust Min and Max to reflect market expectation (possibly from training), or risk preference



Reducing RF

- Desirable to reduce *Rate of Failure*
  - Missing opportunities may be more acceptable than losing money
- Our approach:
  - Augment fitness with constraints
  - Tighter constraints means lower RF
  - Even if lower RF $\Rightarrow$ more missed chances
- Our goal:
  - Allow one to tune RF to one's preference
  - without affecting overall Rate of Correctness

FGP-2 on IBM 1991-1997

- With low rate of failure specified
- Results are more reliable

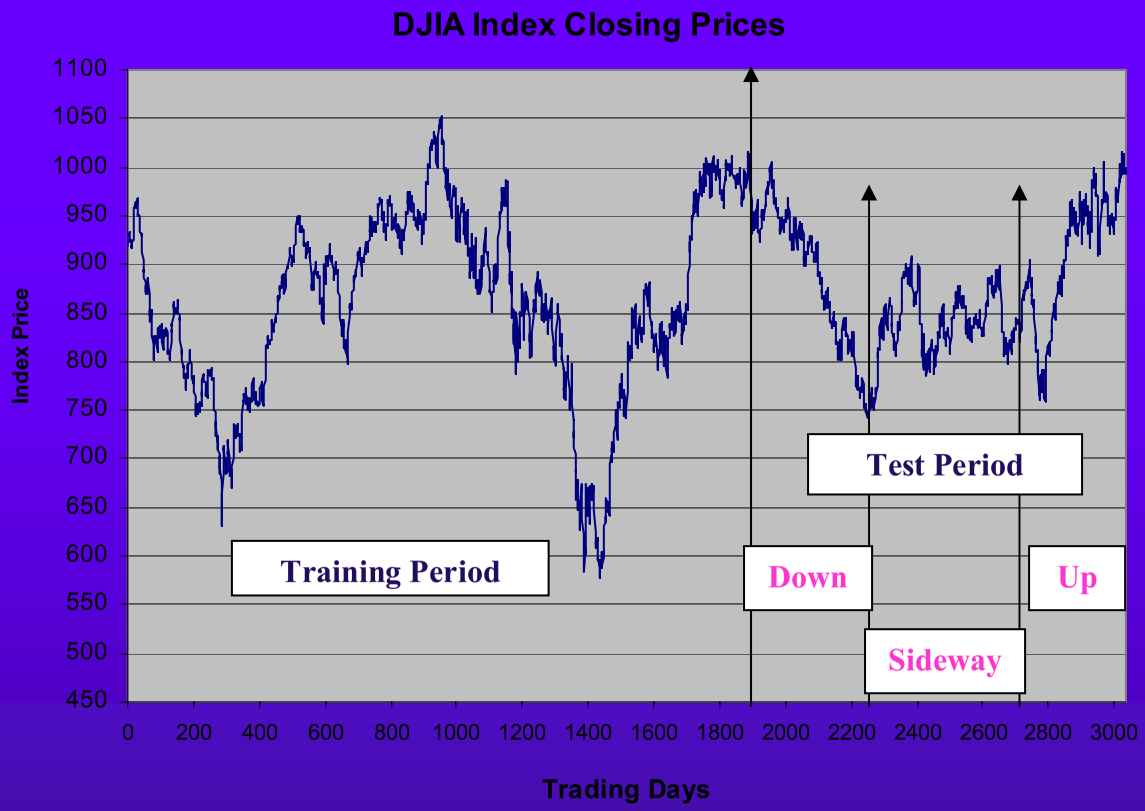
General fitness function:

IBM		Over Test Period (1995.12.28-1997.03.05)			
RC	56.5%	0	1	Opport.:	45.0%
RMC	90.0%	104	6	0	110
RF	40.0%	81	9	1	90
AR	210.0%	185	15	8%	200

Constrained fitness function:

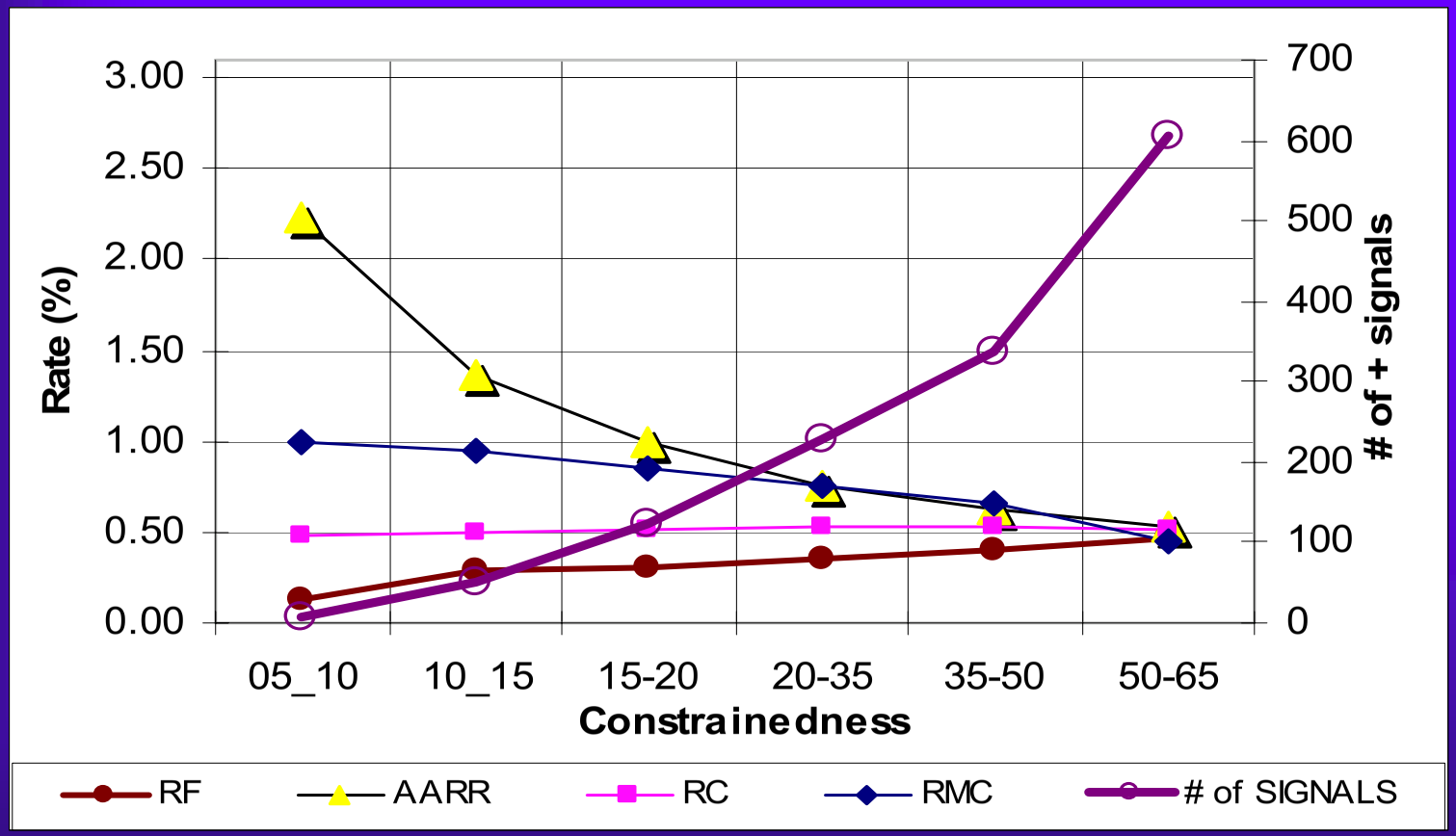
IBM		Test Period (1995.12.28-1997.03.05)			
RC	59.0%	0	1		45.0%
RMC	87.8%	107	3	0	110
RF	21.4%	79	11	1	90
AR	511.0%	186	14	7%	200

FGP-2 on DJIA Data



- Training: 1,900 days 07/04/1969 to 11/10/1976
- Testing: 1,135 days 12/10/1976 to 09/04/1981
- Target: "*rise of 4% within 63 days*"

Effect of constraints on FGP-2



- Observation: RMC can be traded for RF without significantly affecting RC

Our FGP Experience

- Patterns exist
  - Would they repeat themselves in the future?

    (EMH debated for decades)
- EDDIE has found patterns
  - Not in every series

    (we don't need to invest in every index / share)
- EDDIE extending user's capability
  - and give its user an edge over investors of the same caliber

High Frequency Data: Example of an Order Book

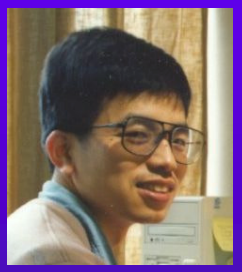
	Price	Volume	Orders
Seller 4	3.86	2,000	1
Seller 3	3.85	10,000	5
Seller 2	3.84	5,000	1
Seller 1	3.83	1,000	1
Buyer 1	3.82	6,000	3
Buyer 2	3.81	8,000	3
Buyer 3	3.80	5,000	1
Buyer 4	3.79	17,000	3

EDDIE in Arbitrage, Historical Note

James Butler

Edward Tsang

1995: EDDIE
(**E**volutionary **D**ynamic **D**ata **I**nvestment **E**valuator)



Jin Li

1996: FGP
(Financial **G**enetic **P**rogramming)

Arbitrage Research

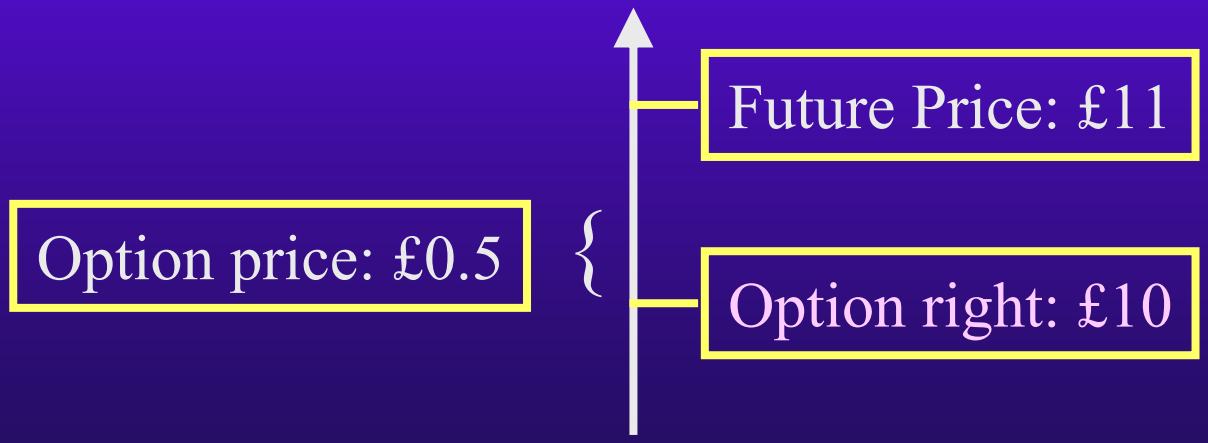
Sheri Markose

Hakan Er

2000:
FGP+Arbitrage

Arbitrage Opportunities

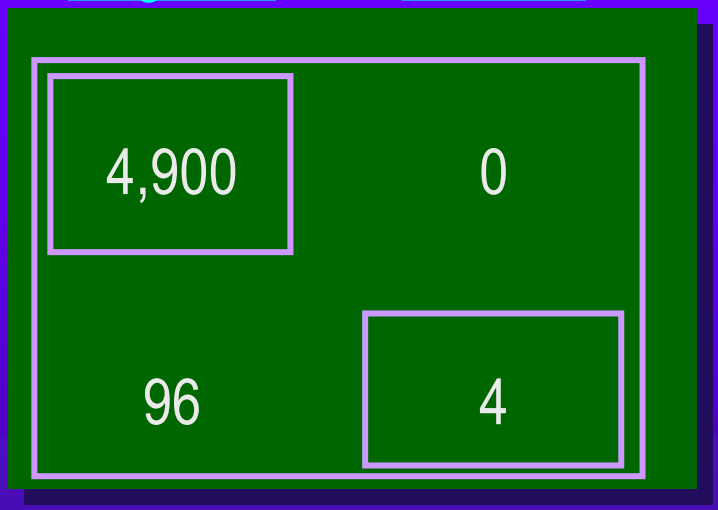
- Futures are obligations to buy or sell at certain prices
- Options are rights to buy at a certain price
- If they are not aligned, one can make risk-free profits
  - Such opportunities should not exist
  - But they do in London



EDDIE in Arbitrage

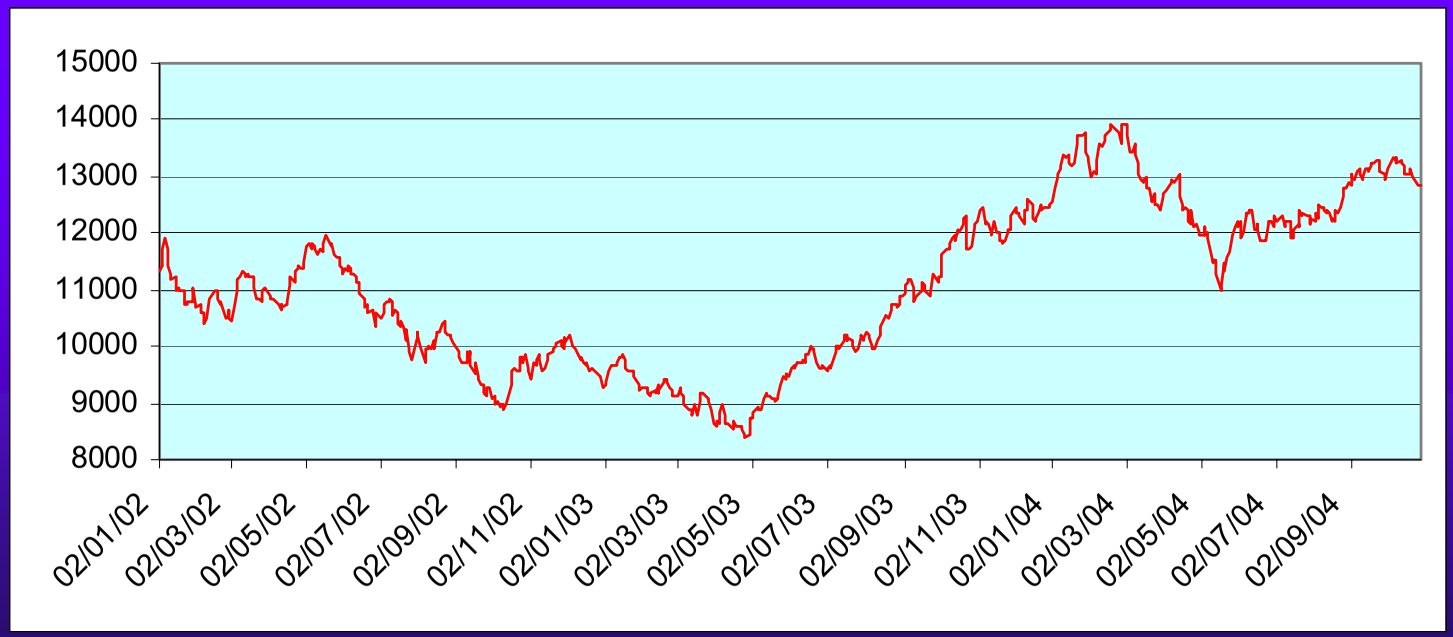
- Arbitrage opportunities found
  - They shouldn't exist?
  - They exist for 12-45 seconds
- EDDIE to predict arbitrage
  - 15 minutes in advance
  - Find clear opportunities only
- Human expertise needed
  - 9 months data preprocessing
  - Over 10 data set revisions

Typical arbitrage result



Profitable arbitrage opportunities are rare; Can't afford to miss too many

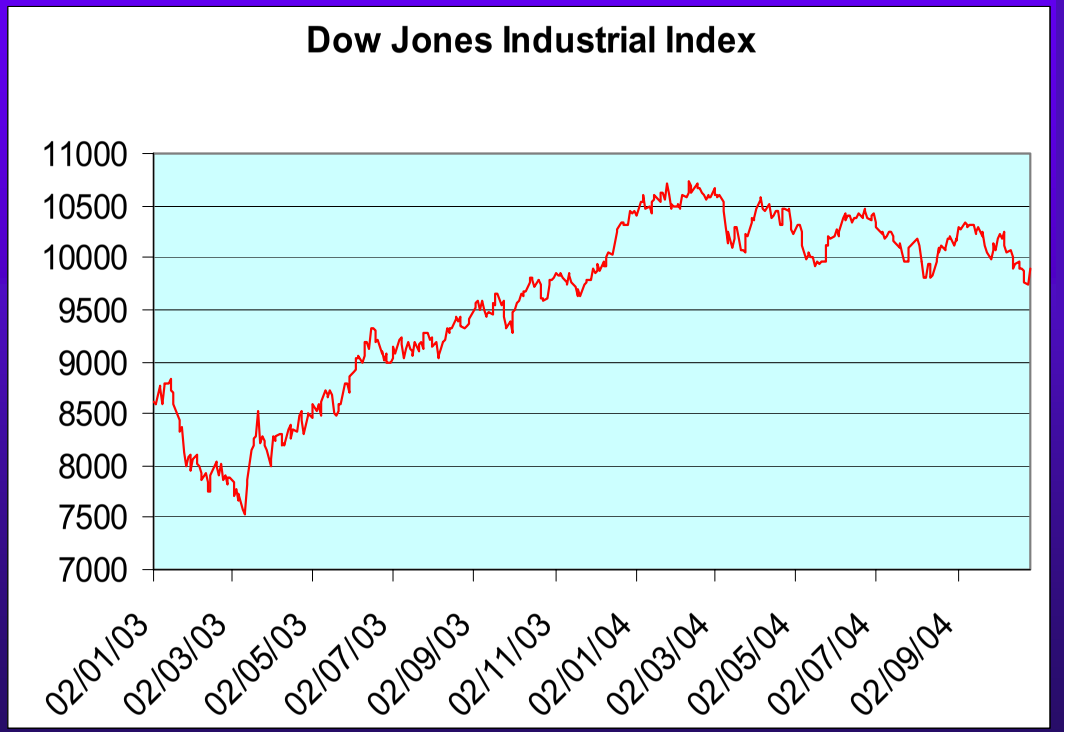
EDDIE Artificial Market



EDDIE as Intelligent Agent

Serafin Martinez
EDDIE Red Q

- Let artificial agents for a market
  - Technical traders (EDDIE)
  - Fundamental traders (Economists)
  - Noise traders
- How would the prices look like?
- Under what conditions will they produce real market stylus?

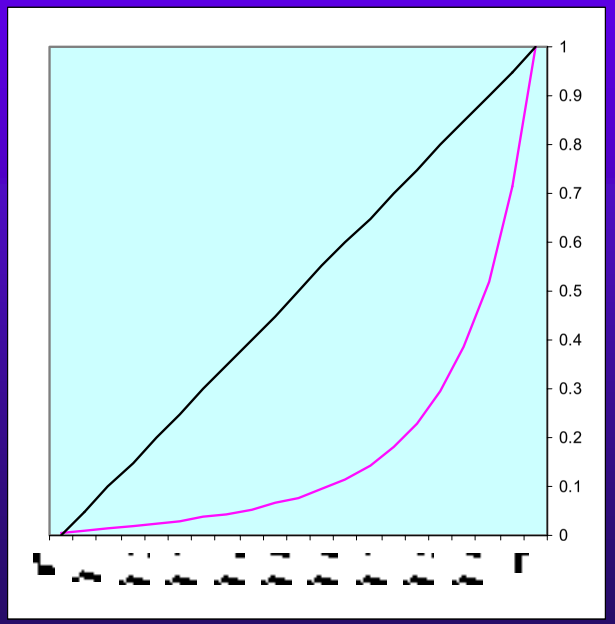


The Red Queen Effect

- Power Law wealth distribution
  - The weakest must re-train themselves
- Red queen effect
  - You have to run as fast as you can to stay in the same place
- EDDIE is used for re-training

Sheri Markose
Red Queen

Serafin Martinez
EDDIE Red Q



EDDIE in Business

From research to practice:

Surfing one step ahead of each wave

29 November 2004

What can EDDIE do for you?

- If it changes 50-50 chances to 55-45
  - in your favour
  - you must be better off in the long run…
- It helps you to beat other investors of the same calibre
  - It provides an extra expert opinion
  - If all your experts give you the same opinion, you have better chance to succeed
- It works day and night, you can't…

EDDIE/FGP is no magic

- A tool is useful when...
  - it can do something good, and
  - we know how to use it, and
  - we know its limitations
- EDDIE / FGP is such a tool
  - No expert input, no useful forecast
    (It only adds value to expert input)
  - It can only find patterns that exist
    (No point asking it to predict the lottery)

Don’t expect to see…

- Miracles –
  - we can’t predict the unpredictable!
- Prediction of everything
  - May not find patterns for the future
    - E.g. patterns found in IBM/BAA, but not HSBC
    - So no positive actions recommended
- Fancy interface
  - At the moment, we concentrate to make EDDIE *predict accurately*

Current Research

- EDDIE for Arbitrage
  - Spot, option and future prices don't always synchronize
  - Hence one can make risk-free return?
- EDDIE for Forecasting
  - When to sell? How to combine signals?
  - What is the return in reality?
- GP for modelling volatility
  - coefficients fitting for GARCH-like functions
  - Discovering new functions forms?
- GP for market understanding
  - Learning agents form artificial market

The Computational Finance Research Team

Edward Tsang
EDDIE / GP

James Butler
EDDIE

Tung L Lau
Early Tools

Jin Li
FGP

Abdel Salhi
Genetic Prog.

Sheri Markose
Red Queen

Hakan Er
Arbitrage

Giulia Iori
Volatility

John Gan
Data mining

Riccardo Poli
Genetic Prog.

Maria Fasli
Agent Tech.

Abhinay Muthoo
Game Theory

Serafin Martinez
EDDIE Red Q

Biliana Alexandrova-Kabadjova
Small world

Nanlin Jin
Bargaining

Alma Garcia
Forecasting

Questions, Discussion