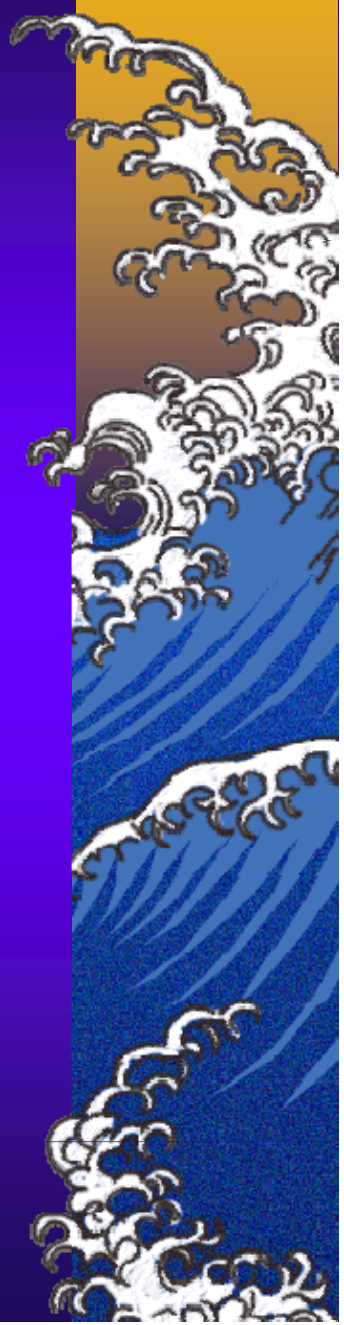
Computational Intelligence Meets Financial Forecasting

Edward Tsang et al

Forecasting Research Team



Computing Intelligence Meets Forecasting

What if opportunities are scarce?

Repository Method

Is the market predictable?

Forecasting System

Predictions, in the form of:
- prices
- opportunities
- threats

EDDIE:
Constraint-directed search
For trading precision with recall

EDDIE for Investment and arbitrage opportunities

How to measure success?

Needs motivate new algorithms

How to invest?

Forecasting

- Will the price go up or down? By how much?

- What data do we have? Daily? Intraday (*high frequency*)? Volume? Indices? Economic Models?
- What is the risk of crashing?
- Are Option and Future prices aligned? (i.e. are there arbitrary opportunities?)

Efficient Market Hypothesis

- Financial assets (e.g. shares) pricing:
  - All available information is fully reflected in current prices
- If EMH holds, forecasting is futile
  - Random walk hypothesis
- Assumptions:
  - Efficient markets (one can buy/sell quickly)
  - Perfect information flow
  - Rational traders

Is the market really efficient?

- Market may be efficient in the long term
- "*Fat Tail*" observation:
  - big changes today often followed by big changes tomorrow (either up or down)
- How fast can one respond to new information?
  - Faster machines certainly help
  - So should faster algorithms (CIDER)
- Credit crunch: did investors price their risks properly?

Do fundamental values matter?

- In boom, markets are liquid but often not driven by fundamentals only (bubbles)
- In bust, markets may be driven by fundamentals only, but are not liquid
- In neither boom nor bust are markets efficient
  - Willem Buiter (LSE)

Our Research agenda

- What would a reasonable agenda be?
- Predicting the price in 10 days would be good
- But it may be sufficient if I could turn a 50-50 game into a 60-40 game in my favour
- Question asked:

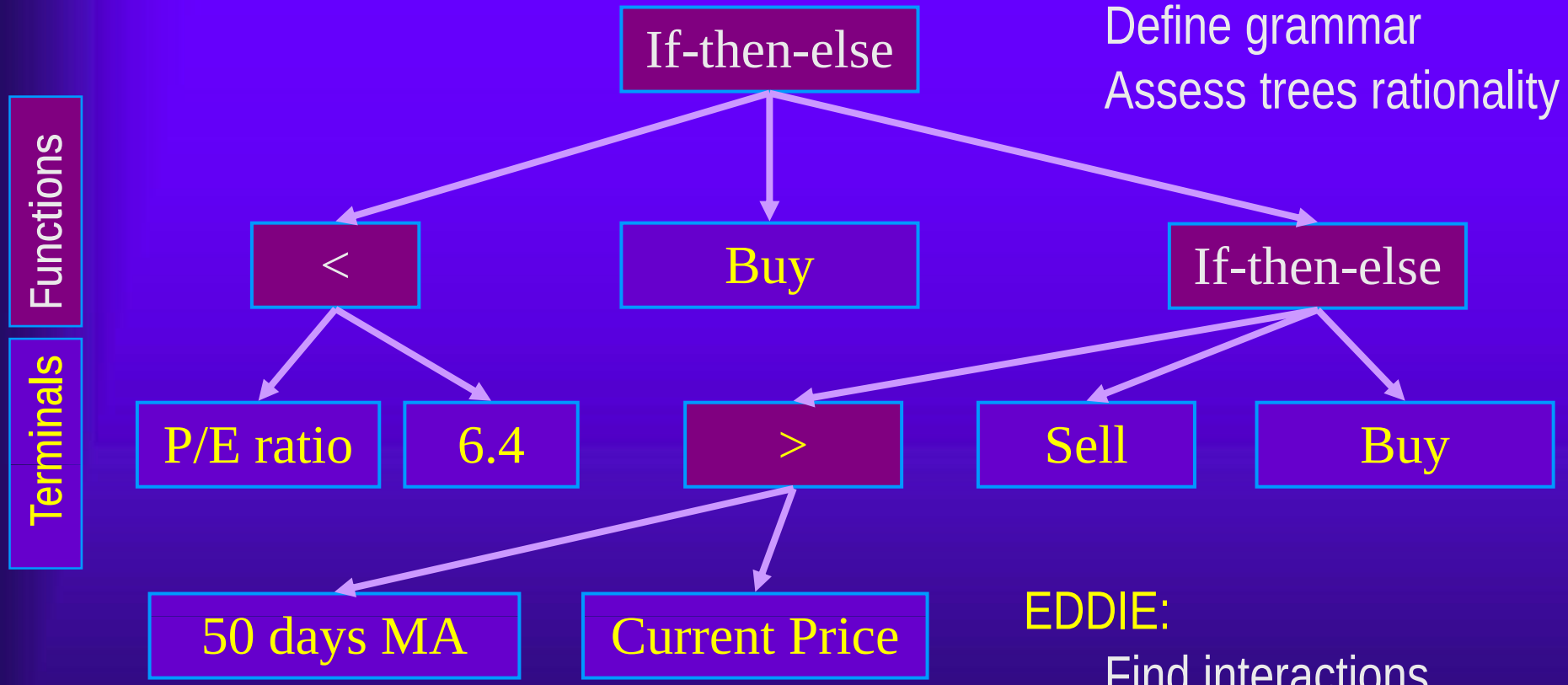
  "*Will the price go up (or down) by at least r% within the next n days?*"

How can computational intelligence help?

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		1
99	82	52		0
87	83	53		1
82	82	51		1
…..	…..	…..		…..

GP: Example Tree



Human users:

- Define grammar
- Assess trees rationality

EDDIE:

- Find interactions
- Discover thresholds

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

Terminals:

Variable is an indicator / feature

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

Threshold is a real number

- Richer language ⇒ larger search space

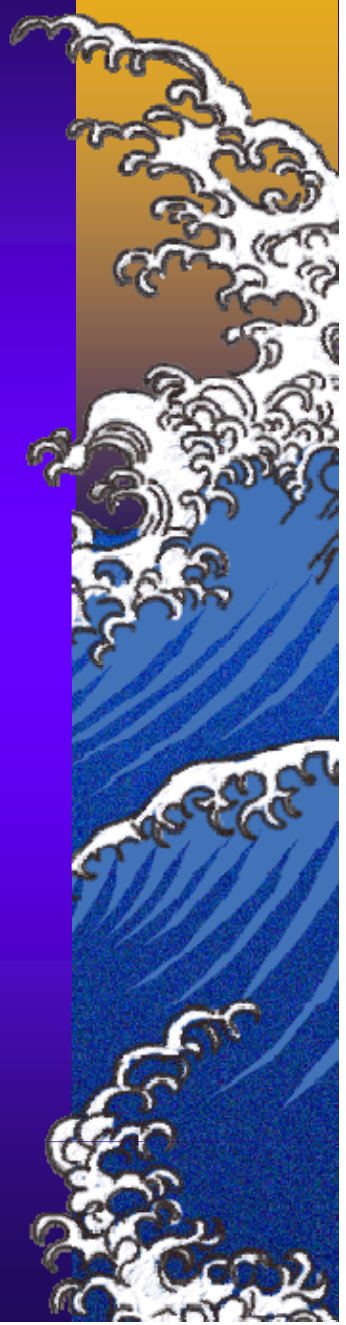
Machine learning basics

What could one learn?

Hypothetical observations

How to summarize success/failure?

Performance measures



Hypothetical Situation

- Suppose you've discovered a good indicator R
  - How can you make use of it?
- Suppose it is a fact that whenever
  - *R* has a value less than 1.4 or greater than 2.7,
  - the volatility of the share prices is above 2.5, and
  - yield is above 5.7%

  prices will rise by ≥6% within the next 21 days
- How can you find this rule

Hypothetical observations

Instance	*R*	Volatility	Yield	Target	Classified	
1	1.2	3.1	4.8	False	False	TN
2	1.3	3.0	6.6	True	True	TP
3	2.8	2.9	5.9	True	False	FP
4	2.5	1.7	7.0	False	False	TN
5	2.4	3.5	6.9	False	False	TN
6	2.0	2.9	5.6	False	False	TN
7	3.1	3.3	5.8	True	True	TP
8	3.1	3.0	5.5	False	True	FN
9	2.8	2.4	5.0	False	True	FN
10	2.6	2.5	5.2	False	False	TN

Confusion Matrix

Reality	Prediction
–	–
+	+
+	–
–	–
–	–
–	–
+	+
–	+
–	+
–	–

Reality	Prediction		
	–	+	
–	5	2	7
+	1	2	3
	6	4	10

Performance Measures

Ideal Predictions

Reality	–	+	
–	7	0	7
+	0	3	3
	7	3	10

Actual Predictions, Example

	–	+	
–	5	2	7
+	1	2	3
	6	4	10

RC = (5+2) ÷10 = 70%

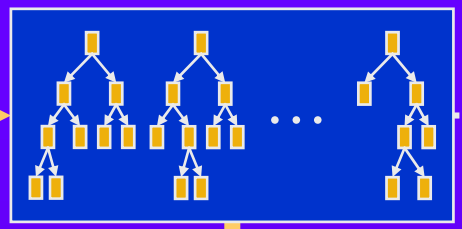
Precision = 2 ÷ 4 = 50%

Recall = 2 ÷ 3 = 67%

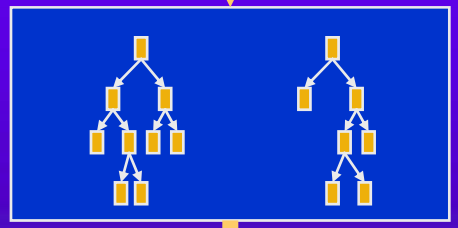
Genetic programming in forecasting

EDDIE

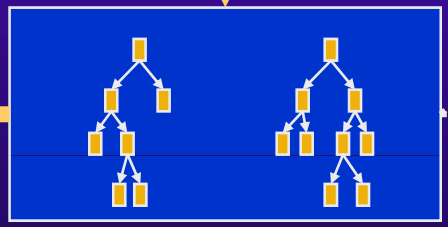
EDDIE Technical Overview



Tournament Selection



Crossover, Mutation



Update population

Tree Representation

Grammar defined by user

Each tree is a Boolean function

E.g. Will the price go up by 4% within the next 21 days?

Constraint-directed fitness

To improve precision

Fitness eval (RC, RF, RMC)

Using historic data

Our EDDIE/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

Incentive to Improve Precision

Actual Predictions, Example

Reality	–	+	
–	75 70	5 10	80 80
+	9 5	11 15	20 20
	84 75	16 25	100

RC = (70+15) ÷ 100 = 85%
Precision = 15 ÷ 25 = 60%
Recall = 15 ÷ 20 = 75%

- False positive costs real money
- We cannot change reality
- But we have control over predictions
- Hope: reduced more false positives than true positive

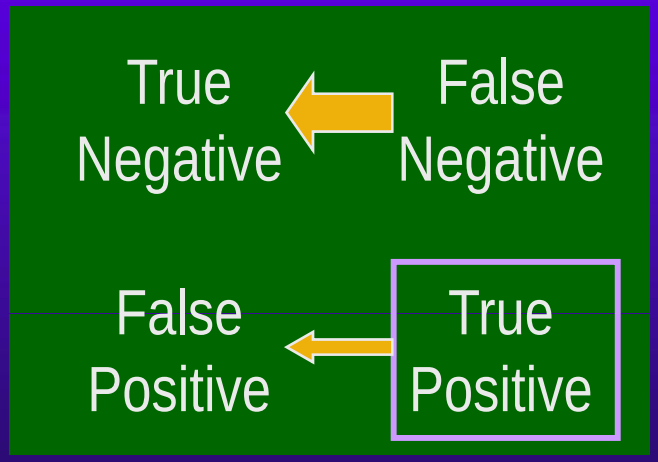
RC = (75+11) ÷ 100 = 86%
Precision = 11 ÷ 16 = 69%
Recall = 11 ÷ 20 = 55%

FGP: Constrained Fitness

Jin Li
FGP

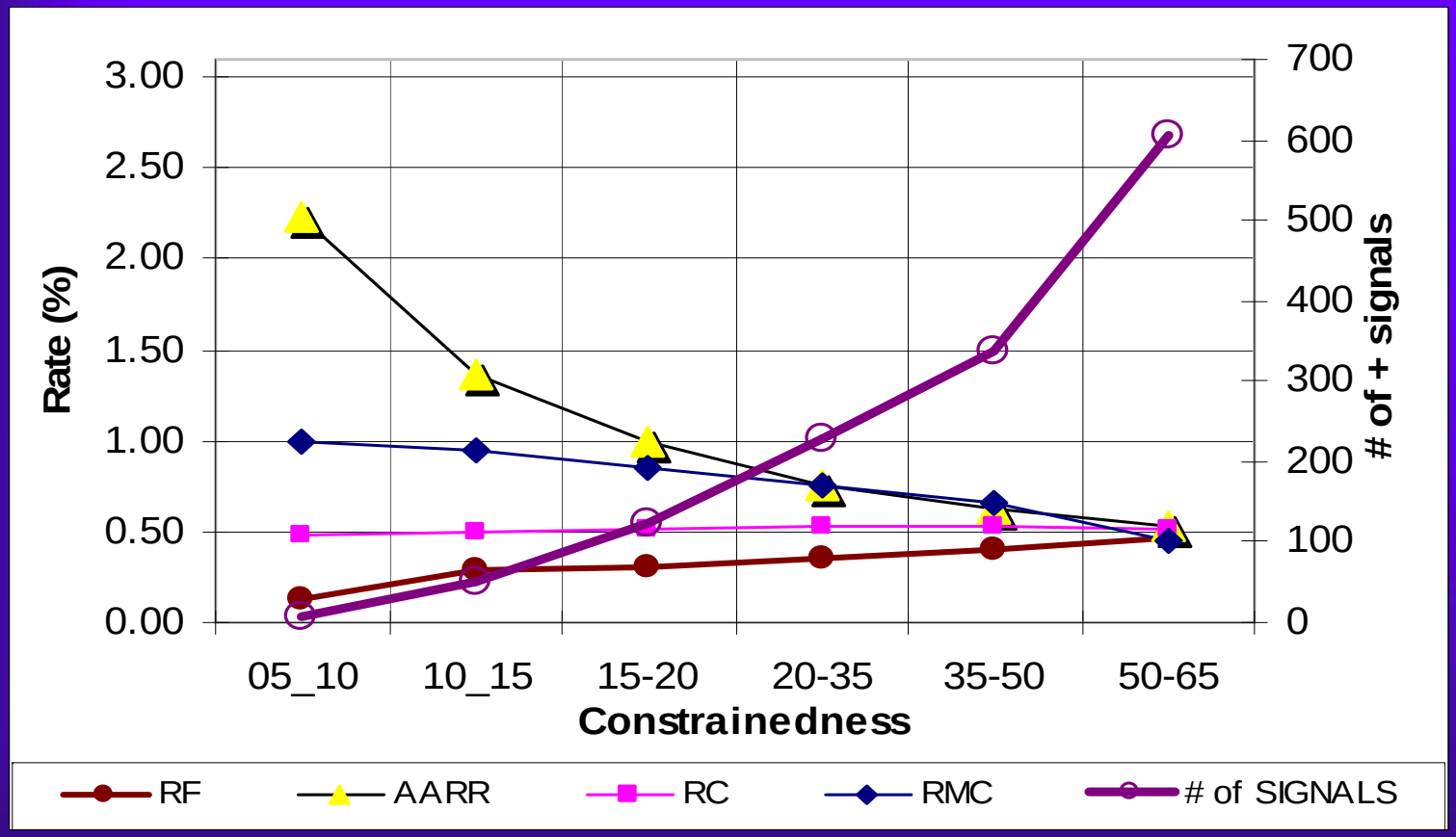
- Constraints can help guiding the search
- Fitness = $w_{rc} \times RC' - w_{rmc} \times RMC - w_{rf} \times RF$
- RC' = RC if P+ $\in$ [Min, Max]
  0 otherwise

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



Cautious ← Low Max

Effect of constraints in FGP-2



DJIA
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*"

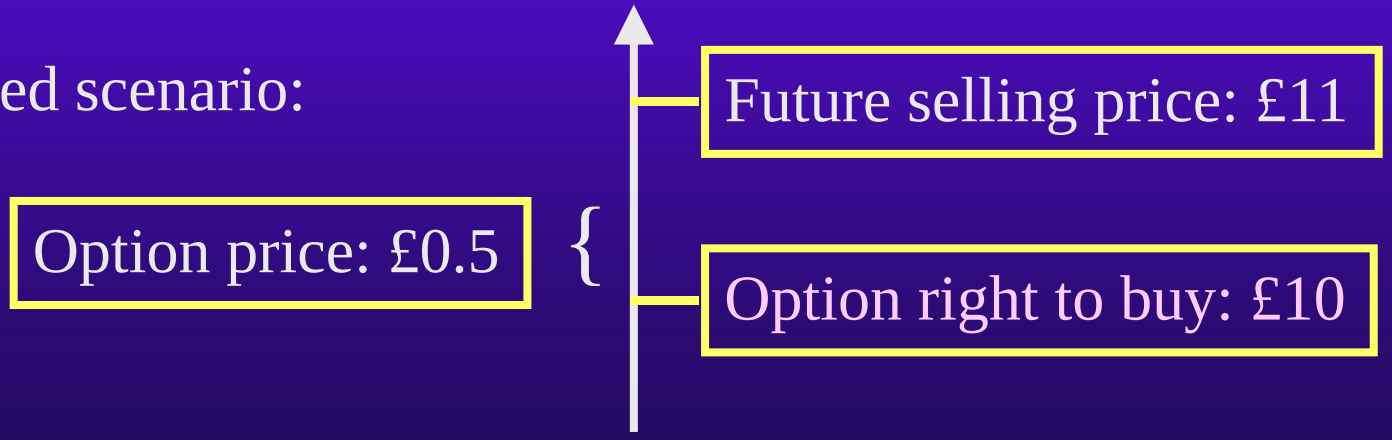
- Observation: RMC can be traded for RF without significantly affecting RC

EDDIE for arbitrage prediction

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

A simplified scenario:



Experience in EDDIE on Arbitrage

- Arbitrage opportunities exist in London
- Naïve approach:
  - Monitor arbitrage opportunities, act when they arise; problem: speed
- Misalignments don't happen instantaneously
  - Do patterns exist? If so, can we recognize them?
- EDDIE-ARB can find some opportunities
  - With high confidence (precision >75%)
- Commercialisation of EDDIE-ARB
  - Need to harvest more opportunities; Need capital
- Research only made possible by close collaboration between computer scientists and economists

Facing scarce opportunities

Chance Discovery

Problem with scarce opportunities

Predictions

Reality	–	+	
–	9,900	0	99%
+	0	100	1%
	99%	1%	

Ideal prediction
Accuracy = Precision = Recall = 100%

Predictions

	–	+	
–	9,801	99	99%
+	99	1	1%
	99%	1%	

Moves from – to +
Random move from – to +
Accuracy = 98.2%
Accuracy = 98.02%
Precision = Recall = 1%
Precision = Recall = 0.01 = 1%
(Accuracy dropped from 99%)

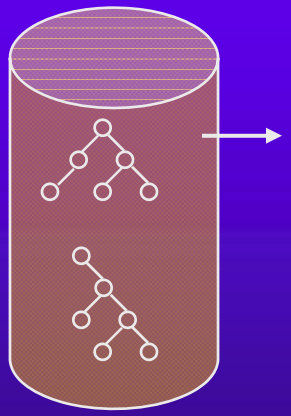
Easy to score on accuracy
Accuracy = 99% – Precision = ?

Repository Method

In order to mine the knowledge acquired by the evolutionary process, Repository Method performs the following steps:

1- Rule extraction

Evolve a GP to create a population of decision trees



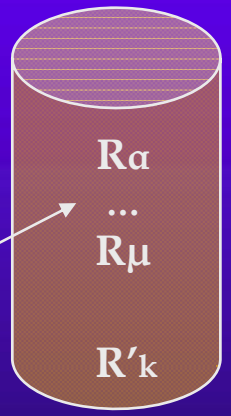
2- Rule simplification

R_1
R_2
...
R_n

The rule R_k is selected by precision;
R_k is simplified to R'_k

3- New rule detection

R'_k is compared to the rules in the repository by similarity (genotype)



4- Add rule to the repository

If R'_k is a new rule R'_k is added to the rule repository

Where does it go from here?

- Computational finance > CI + Finance
  - Research agenda beyond CI and finance experts
- Finance drives computational intelligence
  - We need more techniques for chance discovery
- Being able to forecast alone is not sufficient
  - If opportunity is predicted, do we invest 100%?
- Financial forecasting is growing rapidly
  - Conferences, IEEE Technical Committee, etc

FAQ

FAQ in forecasting

- ***Is the market predictable?***
  - It doesn't have to be
  - But if you believe it is, you should code your own expertise
  - Market is not efficient anyway, herding has patterns
- ***How can you predict exceptional events?***
  - No, we can't
  - Neither can human traders
- ***How can you be sure that your program works?***
  - No, we can't
  - Neither were we sure about Nick Leeson at Barrings
  - Codes are more auditable than humans
  - If you can improve your odds from 50-50 to 60-40 in your favour, you should be happy

Reference

- http://www.bracil.net/finance/papers/Tsang-Forecasting-Fcsc2009.pdf
- Tsang, E.P.K., Forecasting – where computational intelligence meets the stock market, Frontiers of Computer Science in China, Springer, 2009, to appear (also filed as Working Paper WP026-08, Centre for Computational Finance and Economic Agents (CCFEA), University of Essex, revised December 2008)

The Forecasting Research Team

Edward Tsang
EDDIE / GP

James Butler
EDDIE

Jin Li
FGP

Alma Garcia
Chance
Discovery

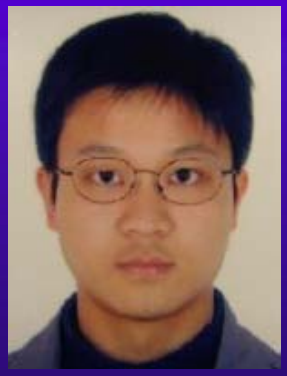
Olsen &
Associates

Sheri Markose
Red Queen

Hakan Er
Arbitrage

Serafin Martinez
EDDIE Red Q

Michael
Kampouridis
EDDIE 8



Wang Pu
EDDIE101

Acknowledgements

Ionic
Share-
scope