Effective Searching for Profitable Forex Trading Rules via Genetic Programming

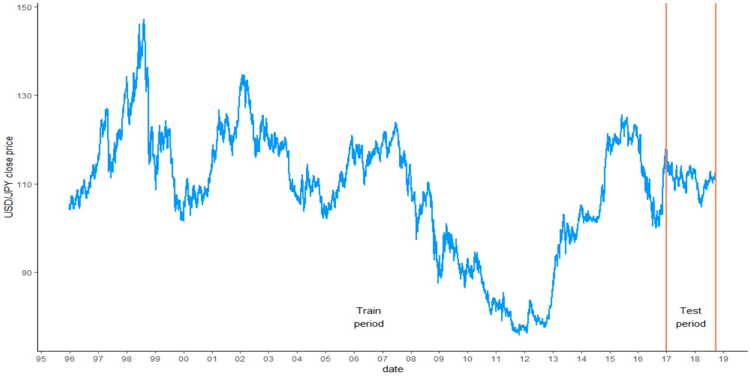
Zelin Chen
Advisor: Dr.Tomasz Woźniak

Department of Economics
University of Melbourne

Oct 2018

Research Question

- Can we let the machine to learn strategies from the exchange rate data (e.g. USD/JPY)?



Genetic Programming

Economics and Econometrics applications

1. Efficient market hypothesis validation
2. Portfolio optimisation
3. Option pricing
4. Risk analysis
5. Causal effects analysis

The Algorithm

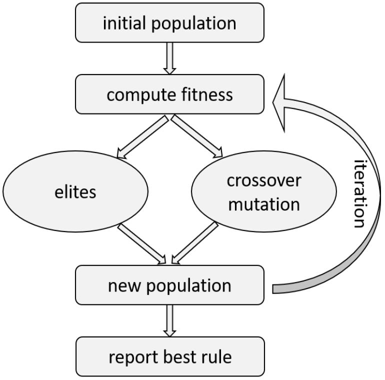


Figure: GP Algorithm(source:[6])

GP building blocks

1. Population
2. Fitness function
3. Evolution functions
   1. Crossover
   2. Mutation
   3. Elitism

Population: Tree-structured trading strategies

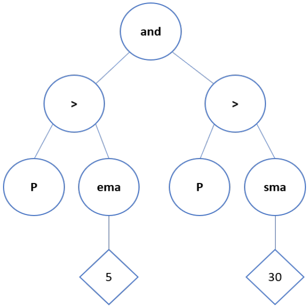


Figure: A long rule sample

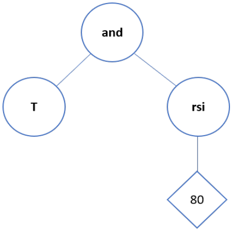


Figure: A short rule sample

Evolution: crossover

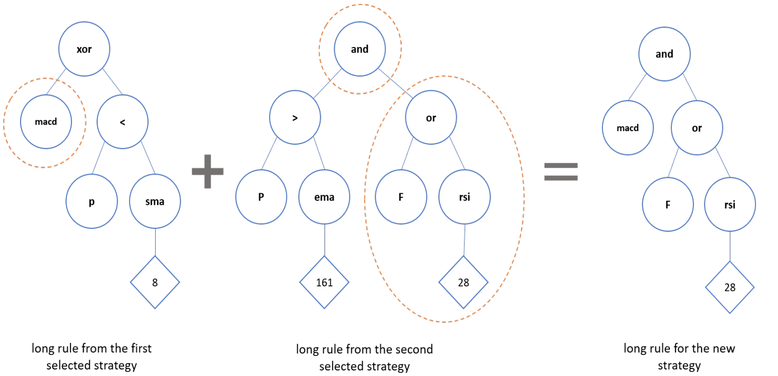


Figure: Crossover for long rules from two selected strategies

Evolution: mutation

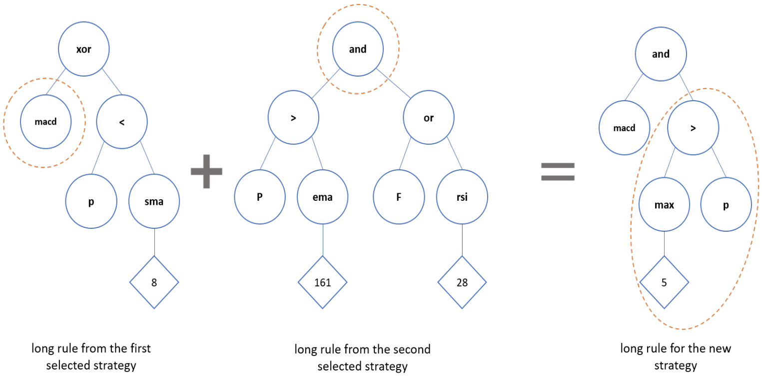


Figure: Mutation for long rules from two selected strategies

GP hyper-parameters

- The hyper-parameters of GP are

hyper-parameter	notation	value
population for each trial	n	30
iteration for each trial	η	20
maximum depth of trees	γ	10
elite ratio	e	10%
elite size	n*e	2
mutation rate	m	5%

- One trial: use GP to iterate one randomly generated population, take the best one in the last generation as the candidate solution. The process is called one trial.
- Repeat the trial for 10,000 times, obtain 10,000 candidate strategies.

Parallel computing in R

- Each trial is independent, i.e. opportunity for parallel computing.
- Given a powerful computer with 32 processors, one can open 30 R at the same time and each R runs for 333 trials.

```
Current generation period:  2 , Max fitness:  110.4113 Time consumption:  2.932164
Current generation period:  3 , Max fitness:  110.4113 Time consumption:  3.498654
Current generation period:  4 , Max fitness:  130.7968 Time consumption:  3.854699
Current generation period:  5 , Max fitness:  130.7968 Time consumption:  4.093063
Current generation period:  6 , Max fitness:  130.7968 Time consumption:  4.077104
Current generation period:  7 , Max fitness:  130.7968 Time consumption:  4.736344
Current generation period:  8 , Max fitness:  130.7968 Time consumption:  5.405556
Current generation period:  9 , Max fitness:  130.7968 Time consumption:  5.817457
Current generation period:  10 , Max fitness:  130.7968 Time consumption:  6.304155
Current generation period:  11 , Max fitness:  170.5584 Time consumption:  5.950101
Current generation period:  12 , Max fitness:  170.5584 Time consumption:  5.867323
Current generation period:  13 , Max fitness:  170.5584 Time consumption:  5.455423
Current generation period:  14 , Max fitness:  170.5584 Time consumption:  5.689796
Current generation period:  15 , Max fitness:  170.5584 Time consumption:  5.404563
Current generation period:  16 , Max fitness:  170.5584 Time consumption:  5.430489
Current generation period:  17 , Max fitness:  170.5584 Time consumption:  5.717721
Current generation period:  18 , Max fitness:  170.5584 Time consumption:  5.832417
Current generation period:  19 , Max fitness:  170.5584 Time consumption:  5.630953
Current generation period:  20 , Max fitness:  170.5584 Time consumption:  5.228031
Trial 352:, Max accumulated principal: 170.558361812492, Time consumption: 1.6706359

-----------------------first stage train-----------------------------------------
Trial 353:
```

A new two-phase searching method

The conjecture

Among 2000 trading strategy candidate, there may be better match-up of long and short rules. For one strategy's long rule, there might be a better short rule match-up from another strategy within 2000 candidates.

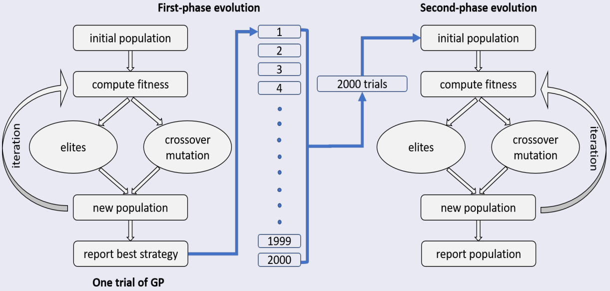
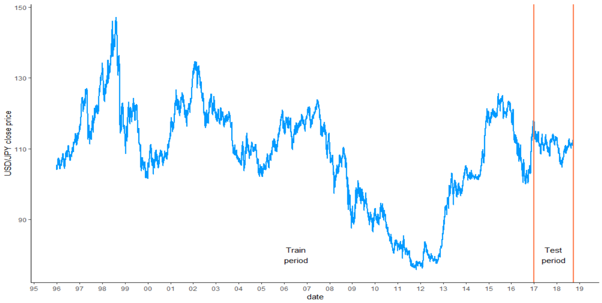


Figure: Two-phase searching method

Forex data



- Use daily JPY/USD exchange rate.
- Train period: 1996-01-01/2016-12-31, test period: 2017-01-01/2018-09-26
- Downloaded from Meta Trader 4, a popular platform among individual traders.

Performance summary

Test period performance of all 10 thousand, top 20, top 5 and the best strategy selected in the training period.

cumulative return (%)	train period		test period	
strategies:	1st phase	2nd phase	1st phase	2nd phase
all 10,000	230.3	506.5	100.3	94.4
top 20	520.3	1119.3	102.6	85.1
top 5	563.7	1122.0	105.0	84.0
top 1	582.0	1126.9	101.6	80.2

Table: Summary of average cumulative returns (%)

Test period performance visualisation

Extensions and Limitations

- Extensions for performance improvement
  1. Include more functions
  2. Fine-tune hyper-parameters
  3. Change data frequency
  4. Increase population size
- Limitation of the research
  1. No optimalisation theory is explained
  2. Searching space is small compares to existing literature.

Franklin Allen and Risto Karjalainen. "Using genetic algorithms to find technical trading rules". In: *Journal of Financial Economics* 51 (1 1999), pp. 245–271.

A. Brabazon, M. O'neill, and I. Dempsey. "An Introduction to Evolutionary Computation in Finance". In: *IEEE Computational Intelligence Magazine* 3.4 (Nov. 2008), pp. 42–55. ISSN: 1556-603X. DOI: 10.1109/MCI.2008.929841.

M. A. H. Dempster and C. M. Jones. "A real-time adaptive trading system using genetic programming". In: *Quantitative Finance* 1 (2001), pp. 397–413.

Alexis Diamond and Jasjeet S. Sekhon. "Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies". In: *The Review of Economics and Statistics* 95.3 (2013), pp. 932–945.

Rafal Drezewski and Leszek Siwik. "Co-Evolutionary Multi-Agent System for Portfolio Optimization". In: 100 (May 2008), pp. 271–299.

John R. Koza. *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. Cambridge, MA, USA: MIT Press, 1992. ISBN: 0-262-11170-5.

C. Lee, C. Quek, and D. Maskell. "A brain-inspired fuzzy neuro-predictor for bank failure analysis". In: *the IEEE International Conference on Evolutionary Computation*. Apr. 2006, pp. 7927–7934.

Christopher Neely, Paul Weller, and Rob Dittmar. "Is technical Analysis in the foreign exchange market profitable? A genetic programming approach". In: *Journal of Financial and Quantitative Analysis* 32 (4 1997), pp. 405–426.

Riccardo Poli, William B. Langdon, and Nicholas Freitag McPhee. *A field guide to genetic programming*. Published via `http://lulu.com` and freely available at `http://www.gp-field-guide.org.uk`, 2008.

F. Streichert, H. Ulmer, and A. Zell. "Evaluating a Hybrid Encoding and Three Crossover Operators on the Constrained Portfolio Selection Problem". In: *CEC 2004*. Apr. 2004, pp. 932–939.

W. Tung and C. Quek. “GenSoOPATS: A brain-inspired dynamically evolving option pricing model and arbitrage system”. In: *CEC 2005*. Apr. 2005, pp. 1722–1729.

Jun Wang. “Trading and Hedging in S&P 500 Spot and Futures Markets Using Genetic Programming””. In: *Journal of Futures Markets* 20.10 (2000), pp. 911–942.

A new two-phase searching method

The conjecture

Among 2000 trading strategy candidate, there may be better match-up of long and short rules. For one strategy's long rule, there might be a better short rule match-up from another strategy within 2000 candidates.

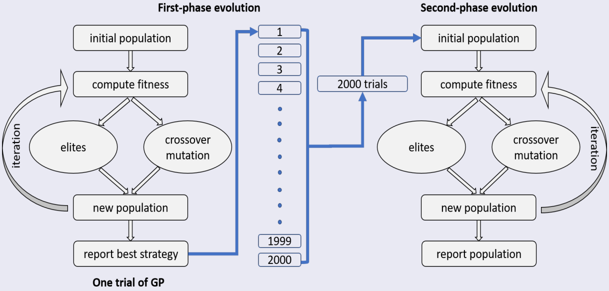


Figure: Two-phase searching method

A new two-phase searching method

- Summary of second-phase GP hyper-parameters

hyper-parameter	notation	value
population size	x	2000
iteration	η	10
elite ratio	e	1%
elite size	x*e	20
mutation rate	m	5%

Table: Summary of GP hyper-parameters in second-phase evolution

Second-phase searching performance boost