Trading Rules on Stock Markets Using Genetic Network Programming with Sarsa Learning

Yan Chen
Graduate school of Information, Production and Systems, Waseda University*
chenyan15153@akane.waseda.jp

Shingo Mabu
Graduate school of Information, Production and Systems, Waseda University
mabu@waseda.jp

Kotaro Hirasawa
Graduate school of Information, Production and Systems, Waseda University
hirasawa@waseda.jp

Jinglu Hu
Graduate school of Information, Production and Systems, Waseda University
jinglu@waseda.jp

Categories and Subject Descriptors

I.2 [**Computing Methodologies**]: Artificial Intelligence

General Terms

Algorithms and Economics

Keywords

Genetic Network Programming, Reinforcement Learning, Sarsa, Stock Trading Model, Technical Index, Candlestick Chart

In this paper, the Genetic Network Programming (GNP) for creating trading rules on stocks is described. GNP is an evolutionary computation, which represents its solutions using graph structures. It has been clarified that GNP works well especially in dynamic environments since GNP can create quite compact programs and has an implicit memory function. In this paper, GNP is applied to creating a stock trading model. There are three important points: The first important point is to combine GNP with Sarsa Learning which is one of the reinforcement learning algorithms. The second important point is that GNP uses candlestick chart and selects appropriate technical indices to judge the buying and selling timing of stocks. The third important point is that sub-nodes are used in each node to determine appropriate actions (buying/selling) and to select appropriate stock price information depending on the situation. In the simulations, the trading model is trained using the stock prices of 16 brands in 2001, 2002 and 2003. Then the generalization ability is tested using the stock prices in 2004. From the simulation results, it is clarified that the trading rules of the proposed method obtain much higher profits than Buy&Hold method and its effectiveness has been confirmed.

GNP-Sarsa consists of judgment nodes and processing nodes, which are connected to each other. Judgment nodes have if-then type branch decision functions. They return judgment results for assigned inputs and determine the next node. Processing nodes take actions (buying or selling stocks). The role of a start node is to determine the first node to be executed. GNP-Sarsa has two kinds of time delays: time delays GNP-Sarsa spend on judgment or processing, and the ones it spends on node transitions. In this paper, important indices and candlestick chart are used as judgment functions, while a_{ip} is used as a threshold for determining buying or selling stocks in a processing node. The role of time delays is to determine the maximum number of technical indices and candlestick information to be considered when GNP-Sarsa determines buying or selling at a certain day.

In the learning phase, Sarsa can obtain Q values which estimate the sum of the discounted rewards obtained in the future. Suppose an agent selects an action a_t at state s_t at time t, a reward r_t is obtained and an action a_{t+1} is taken at the next state s_{t+1}. Then $Q(s_t, a_t)$ is updated as follows.

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha\, [r_t + \gamma Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)]$$

α is a step size parameter, and γ is a discount rate which determines the present value of future rewards: a reward received k time steps later is worth only γ^{k-1} times of the reward supposed to receive at the current step.

In the evolution phase, the role of evolution is to change graph structures and randomly change node parameters a_{ip}.

To confirm the effectiveness of GNP-Sarsa, we carried out the trading simulations using 16 brands selected from the companies listed in the first section of Tokyo stock market in Japan. From the simulation results, it is clarified that in the training term the fitness becomes larger as the generation goes on and the profits obtained in the testing term are better than Buy&Hold in the simulations of 14 brands out of 16 and even when there is downtrend, although Buy&Hold makes a loss in five brands, the proposed method can obtain profits in four brands and also decrease the loss in one brand.

For confirming the optimization function of GNP-Sarsa, we also calculate the average ratio of the nodes used in the test period over 10 independent simulations in order to see which nodes are used and which are most efficient for stock trading model. From the result, we can see that the judgment nodes of ROC and RCI are frequently used. Thus it can be said that GNP-Sarsa judges that these nodes are important to determine stock trading. GNP-Sarsa can automatically determine which nodes should be used in the current situation by evolving node functions and connections between nodes, in other words, GNP-Sarsa can optimize the combination of technical indices and candlestick chart used for stock trading model.

GECCO'07, July 7–11, 2007, London, England, United Kingdom
ACM 978-1-59593-697-4/07/0007.

* Hibikino 2-7, Wakamatsu-ku, Kitakyushu, Fukuoka 808-0135, Japan. Tel/Fax: +81 93 692-5261.