STOCK TRADING USING XCS

by

SRINIVAS PEDDOLA

B.E., Osmania University, 2008

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Computing and Information Sciences
College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2011

Approved by:

Major Professor
Dr. William H. Hsu

Abstract

Stock trading is a complex process which is subject to distinct events both inside and outside an organization. An increase in revenue is a direct influence which causes the stock price to move upwards; likewise, the price of a stock fluctuates due to indirect influences. For example, a stock's price may move upwards due to a firm arriving at beneficial deals or adding eminent professionals to its board of directors; a rise in correlated stock prices. The stock prices are affected to a great extent by statements of the finance minister and other related officials. In addition, subjective judgments and emotions of traders can also influence the variation of indices and stock prices in the market.

The efficient market hypothesis proposes that stock price is unpredictable, assuming all past information has been influenced on current price and therefore it is not useful for the prediction of future price. Nevertheless, there are opposing theories which state that stock prices are predictable through the identification of trends and price patterns based upon past data such as price and volume quotes, balance sheets, and income statements.

In the stock market, naive traders (or investors) assume risks due to the above uncertainties, but still have opportunities to make profits through proper, in-depth analysis on sufficient quantities of past data. There are many indicators that are accessible and can help predict the direction of future prices or index values using fundamental and technical data. Fundamental data, derived from the balance sheets and income statements, is preferred for mid-term and long-term investors but not suitable for short-term investors; meanwhile, technical data can be used for short-term investors as well. Technical data is preferred for short-term investments but it can also be used for long-term investments by choosing a specific window of time to look back in determining the indicators for long periods.

The current trading model for stocks and indices was developed using an accuracy-based learning classifier system (XCS), which combines reinforcement learning, genetic algorithms, and other heuristics to form an adaptive system whose purpose is to execute stock trades for profit. A test bed developed for experimenting with this system consists of technical data, with candidate features chosen as the most popular indicators.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I owe my deep gratitude towards my academic advisor and major professor, Dr. William Hsu, for providing his valuable suggestions and guidance throughout the masters program. It has been a rewarding experience for me to work with him. I feel privileged to have joined his Knowledge Discovery in Databases research group.

I am grateful to Dr. Doina Caragea and Dr. Torben Amtoft for graciously accepting to serve on my committee. I would like to thank them for reviewing my thesis and offering their important feedback and suggestions.

I would like to thank Mr. Prabhakar Thennarasu for sharing his knowledge on stock trading using technical analysis which helped me in understanding factors that affect the stock market, and the consequences of gambling in the stock market.

I would also like to express a special thanks to my family members for their blessings, good wishes, undying love and support throughout my studies.

Chapter 1 - Introduction

Analyzing and forecasting stock prices or index changes has been the subject of a great deal of research in the field of machine learning and pattern recognition. Many stock investors prefer to depend on trading agents that assist in predicting the direction of future prices based on various situations and thus help in making instantaneous decisions. This is because it is established that the ability to forecast the direction and not the exact value of future stock prices is the most significant factor in earning wealth through financial prediction (Choudhry & Garg, 2008).

Stock prices are highly dynamic and easily affected by rapid changes because of the basic nature of the financial domain, and also due to the combination of known parameters such as the previous closing price, change in volume and unknown factors such as election results, financial statements, rumors (Tsaih, 1998). Thus, a stock trading agent must predict the direction of future prices over a specified term, and recommend a course of action to the investor, such as bidding on a stock before its price rises, or selling it before its value declines (Tsaih, 1998). The trading agent must be capable of learning in a sense that its ability to predict must improve with experience through feedback from the stock market simulator and evolving accurate rules. An accurate prediction algorithm would turn out higher returns for the investor indicating a direct relationship between the accuracy of prediction algorithm and the amount of profits earned using the learning classifier system or any trading model (Shah, 2007).

Stock market data can be subdivided into three major categories: technical data, fundamental data and derived data (Hellstrom & Holmstrom, 1998). Technical data contains all the information that is related to stocks. It includes historical quotes of a corporation that has volume and price data. Fundamental data contains all the information that is referred to the intrinsic value of a corporation and the data related to stock market (Hellstrom & Holmstrom, 1998). It includes interest rates, trade balance, prices of related commodities, inflation, net profit margin of a firm, prognoses of future profits, and prognoses of future sales. (Hellstrom & Holmstrom, 1998). The derived data contains a kind of information that can be produced through transforming and combining technical and/or fundamental data (Hellstrom & Holmstrom, 1998).

1.1 Goal

The objective is to implement an accuracy-based learning classifier system (XCS) for the stock market simulator using the Java programming language and evaluate its results.

The XCS stock trading model is evaluated using a payoff-based mechanism. The total payoff is a sum of numeric rewards received from an environment after the prediction of output for each input supplied to a trading agent. The numeric value or payoff represents an amount of profits or losses incurred upon executing a proposed action where a positive value of payoff (or pleasure) indicates a profit, a negative value of payoff (or pain) indicates a loss, and a zero is an indication of neither profit nor loss.

1.2 Motivation

The prediction of stock price (or index) is a challenging sub problem because there are some theories such as efficient market hypothesis that assume stock prices (or indices) are completely unpredictable. It is because the entire stock information, such as fundamental data and technical data, already gets influenced on the current stock price (or index). Thus forecasting of stock price (or index) using the available information makes no sense and thus become useless. There are proponents on also the alternative side who believe that stock prices (or indices) follow past trends, patterns and could be predicted using the indicators that are derived from fundamental data and technical data.

The stock markets play a major role on the economic situation of a nation. The gross domestic product of a nation as well is dependent on the countries stock exchange index value. Traders (or investors) could easily make a lot of wealth with the accurate prediction algorithms. There are numerous trading tools available over the internet while most of those use just the technical analysis, some of those even use fundamental analysis but there were not enough tools based on machine learning algorithms.

This thesis will propose an implementation of learning classifiers systems to the stock market simulator with an aim to earn optimal returns.

1.3 Prediction Methods

In practice there are four prediction strategies which are technical analysis, fundamental analysis, traditional time series and machine learning methods (Kalyvas, 2001).

1.4 Fundamental Analysis (FA)

Fundamental analysis is concerned more with corporation rather than actual stock (Shah, 2007). The aim of this prediction method is to estimate the actual value of stock to invest and determine its real value through examining the parameters such as growth, dividend payout, interest rates, risk of investment, sales level, tax rates, and so on (stockcharts.com, 2011). FA is preferred for long-term investments because the fundamental data such as income statement, dividend payout, balance sheet of a corporation will be released just a few times within a financial year. Fundamental analysts believe that stock market is affected with the logical factors rather than psychological factors (stockcharts.com, 2011). FA is not affected with the psychological factors because it is not derived from historical quotes, price patterns, and technical data that keep on changing all the time per a fraction of second. The trading agents could perform much better on combining the technical analysis with the fundamental analysis which is the belief of most financial experts. We have not employed fundamental analysis in our data set as it does not fit in the objectives of our research work.

Fundamental analysis has both strengths and weaknesses; Strengths are that it makes use of long-term trends, value spotting, business acumen, and knowing who is who (stockcharts.com, 2011). Among its weaknesses are brittleness due to overreliance on industry-specific (or company-specific) indicators such as time constraints, subjectivity, analyst bias, and definition of fair value (stockcharts.com, 2011).

Indicators derived from fundamental data that focus on assets, earnings, dividends, and sales in market are used to forecast the orientation of future price. There are numerous fundamental indicators that assist in making predictions but no single indicator is sufficient to generate a signal to either bid or offer a stock. However these are the benchmarks to measure the worthiness of potential investments. The description and formulas of these indicators are not shown in our thesis as the data set supplied to the trading model contains technical indicators alone.

Some of the examples for fundamental indicators are earnings per share, price to earnings ratio, return on equity ratio, working capital ratio, projected earnings growth, dividend payout ratio, dividend yield, and book value.

1.5 Technical Analysis (TA)

Technical analysis deals with the affect of price based on trends, patterns and other indicators using a time series analysis. The idea behind technical analysis is that prices move in predictable trends which are driven through the invariably changing attributes of investors in response to different forces. A technical analyst use charts to forecast the direction of future price movement applying the past technical information such as open and close prices, volume of shares, highest and lowest prices during a trade period. Technicians consider the market to be 80 percent psychological and 20 percent logical (stockcharts.com, 2011). In most of the cases price charts are used to recognize trends that are assumed to be based on supply and demand issues that often have noticeable patterns (stockcharts.com, 2011).

1.5.1 Trend Lines

A trend line is a straight line that is drawn through connecting two or more peak points that extends into future to represent a support and resistance levels which plays an important role in technical analysis to identify and confirm a trend (stockcharts.com, 2011). Most of the rules that were applicable to support and resistance level could also be applied to the trend lines (stockcharts.com, 2011).

The three different forms of trends are an uptrend with the increasing value in price, a down trend with the decreasing value in price, and a trading range where the price oscillates in between the upper and lower price levels.

Uptrend Line

Uptrend signifies a rise in stock price or index value. It has a positive slope which is formed on joining atleast two lower peak points; the second lower peak point must be higher than the first lower peak point in order to have a positive slope (stockcharts.com, 2011). An uptrend line acts like a support which is an indication that the demand is increasing despite the rise in stock price

(Stock Charts). As long as the stock price remains above a trend line uptrend is solid and intact which is an indication of good sign to invest (stockcharts.com, 2011).

The following chart which is shown here is an instance of uptrend in stock price followed by the EMC Corporation during July 1997 to January 2001 for a span of about 42 months. The horizontal axis represents a time frame and the vertical axis represents a closing price in this scenario. The points marked on price chart are support levels at different time frames.

Figure 1 - Uptrend Line

Downtrend Line

Downtrend signifies a fall in stock price or index value. It has a negative slope which is formed on joining atleast two higher peak points; the second higher peak point must be lower than the first higher peak point in order to have a negative slope (stockcharts.com, 2011). A downtrend acts like a resistance which is an indication of decrease in demand with the rise in stock price (Stock Charts). As long as the stock price remains below a trend line the downtrend is regarded as solid and intact which is an indication of good sign to sell the stock (stockcharts.com, 2011).

The following chart shown here is an instance of downtrend in stock price followed by the Amazon.com, INC during December 1999 to December 2001 for a span of about 24 months. The horizontal axis represents a time frame and the vertical axis represents a closing price in this scenario. The points marked on price chart are resistance levels at different time frames.

Figure 2 - Down Trend Line

Trading Range

Trading range is a time frame during which the stock price or the index oscillates within a relatively tight range which is an indication of equally balanced demand and supply forces on the stock or index (stockcharts.com, 2011). A break down above the trading range is a win for bulls, investors those who make profits when the price moves upwards, and a breakdown below the trading range is a win for bears, investors those who make profits when the price moves downwards (stockcharts.com, 2011). It is useful in identifying the supports and resistance levels as turning points or continuation patterns (stockcharts.com, 2011).

The following price chart is a trading range example of Halliburton from December, 1999 to March 2000. The horizontal axis represents a time frame and the vertical axis represents a closing price in this scenario. The points marked on chart are support and resistance levels.

Figure 3 - Trading Range

1.5.2 Support and Resistance

Support and *resistance* are the price levels at which the movement of stock price should stop and reverse its direction thereafter. “Support and resistance represent key junctures where the forces of supply and demand meet. In the financial markets, prices are driven by excessive supply (down) and demand (up). Supply is synonymous with bearish, bears and selling. Demand is synonymous with bullish, bulls and buying. These terms are used interchangeably throughout this and other articles. As demand increases, prices advance and as supply increases, prices decline. When both supply and demand are equal, prices move sideways as bulls and bears slug it out for control.” (stockcharts.com, 2011)

In the above price chart Figure 3 - Trading Range red line illustrates a reistance level and green line illustrates a support level. The investors will opt to invest on stocks when closing price is near the support level and will opt to sell the stocks when closing price of a stock is near the resistance level.

Whenver the stock price crosses resistance level the former resistance level would become a new support level which is a good sign to invest on stock at this situation. Similarly, Whenver the stock price crosses support level the former support level would become a new resistance level which is an indication to sell stock at this situation. Thus a breakdown of resistance level is a win of bulls and a price below the support level is a win of bears.

1.5.3 Price Patterns

"Price patterns are simply more complex versions of trend lines". (stockcharts.com, 2011)

Price pattern analysis is used for both short-term and long-term investments. The stock price data could be intraday, daily, weekly or monthly and the price patterns are usually as short as a day or as long as many years (stockcharts.com, 2011). The time frame used for intraday trading is usually 1-minute, 5-minute, 10-minute, 15-minutes, 30-minutes or hourly which is nothing but the price variations in between the opening price, closing price during a trading period and for end of day (EOD) trading the time frame could be daily, weekly, or monthly and the stock price used for analysis is the closing price. There are various price patterns that are used for technical analysis, some of the most famous and reliable chart patterns are head and shoulders, cup and handle, rounding patterns, double top, double bottom, triple top, triple bottom, ascending triangle, descending triangle, and symmetrical triangle.

1.5.4 Technical Indicators

The purpose of technical indicator is to anticipate in advance whether the future price is going to rise or fall and thus suggests the trader to either bid a share or offer a share. Technical indicators recommend the investors to sell a stock if the forecast is fall in future price or invest on a stock if the forecast is rise in future price. Technical indicators are derived from mathematical formulae using historical price and volume data of a stock. The price data includes combination of open, high, low, close over a specific period of time. Most of the indicators are derived from closing price while others include both volume and price data into their formulas. Indicators, such as moving averages, are derived from simple formulas and the mechanics are somewhat simple to understand (stockcharts.com, 2011). Other indicators, such as RSI, have complex formulas and require more study to understand and appreciate (stockcharts.com, 2011).

No single indicator is sufficient for the analysis, so a combination of multiple indicators is used to generate a signal (or action) to execute. A collection of indicators that give higher returns for one stock might not work in a similar manner for other stocks. It is real difficult to come up with a combination of indicators suitable for a particular stock. So we are using a collection of nine indicators that are used in wide and also the most popular indicators used for the stock trading. In our data preparation for the XCS trading agent the following technical indicators are used that makes use of entire technical information available.

Moving average (SMA and EMA)

"Moving averages smooth the price data to form a trend following indicator. They do not predict price direction, but rather define the current direction with a lag. Moving averages lag because they are based on past prices. Despite this lag, moving averages help smooth price action and filter out the noise. Moving averages are the basic building blocks for many other technical indicators and overlays such as Bollinger bands and MACD." (stockcharts.com, 2011)

Moving average is measured on computing the average of prices over past n-periods. For intraday trading purpose the moving average is measured by taking the average of past prices where the time scale would be in minutes where as for the end of day trading the time frame could be in days, weeks or months depending on the investment period.

The number of periods to use depends on whether the trader is interested in short-term investment or long-term investments. Smaller moving averages are used for the forecasting of short-term trading and longer moving averages are used for the forecasting of long-term trading.

Usually, a 10-minute time frame is used for the intraday trading and a single day time frame is used for the end of day trading purpose.

The estimate of moving average of a stock is skipped for the initial n-periods and the first value of moving average is computed at nth period by taking the average of prices over the past n-periods assuming the data in incremental order of time series. The moving average of next period is computed by dropping the first price value and adding the value of price in the next period and so on. Thus the indicator acts like an average of prices over a moving window of n-time periods that moves a single period after every calculation of moving average from top to bottom.

There are two important types of moving averages which are simple moving average and exponential moving average. Both use the same concept of taking an average of past prices where the simple moving average gives equal weight to all prices however the exponential moving average gives an importance to the most recent prices compared to the old price values.

Formulae:

The formulas used to measure the indicators are adopted from a website www.stockcharts.com

$$\text{n period SMA} = \text{Sum of N periods price} \div \text{n}$$

$$\text{Multiplier} = 2 \div (\text{n} + 1)$$

$$\text{n period EMA} = \text{Closing Price - EMA(previous period)} \times \text{multiplier} + \text{EMA(previous period)}$$

Where n is the number of periods used for the calculation of moving averages.

Simple moving averages are used useful in identifying the potential support and resistance price levels where as exponential moving average that lags a lesser extent and more sensitive to recent price changes will turn before the simple moving averages when plotted. Therefore it is useful in identifying a prior change in trend.

Moving average convergence divergence (MACD)

MACD is developed by Gerald Appel in late seventies which turns two trend following indicators into a momentum oscillator which is measured by taking the absolute difference between shorter exponential moving average and longer exponential moving average (stockcharts.com, 2011).

The reason for choosing exponential moving average over simple moving average for estimating MACD is because exponential moving average lags to a lesser extent than simple moving average as it gives higher weights to the recent prices which is helpful in identifying a prior change in trend.

Formula:

$$\text{MACD} = \text{m period EMA - n period EMA} \quad (m < n)$$

The shorter moving average lags lower than the higher moving average. Indicator is positive when the shorter moving average value is higher than the longer moving average and negative when the shorter moving average is lower than the longer moving average.

The plot of an indicator oscillates above and below the zero line and the zero value of an indicator represents a crossover. The crossover is an indication of change in trend that happens when both the moving averages are equal in value. The crossover of shorter moving average curve from above the longer moving average curve then it is a prior signal of fall in price in near future. Similarly, the crossover of shorter moving average curve from below the longer moving average curve then it is a prior signal of rise in price in near future.

The traders look for convergences which is an indication of centerline crossover and divergences in order to generate buy and sell signals. The indicator is not useful for the identification of overbought and oversold levels as the values of this momentum oscillator are unbounded (stockcharts.com, 2011). The range of indicator depends on the value of price as it just the absolute difference, the stocks with higher price value would have higher range and vice versa.

The default settings used for the estimation of MACD are 12, 26 and 9 which is an absolute difference of 12 periods moving average and 26 periods moving average with the 9 periods moving average as a signal line (stockcharts.com, 2011). The traders also use signal line crossovers along with the center line crossovers to forecast the future price. The plot of indicator and the signal line is used to identify the crossovers where the crossover of indicator curve from below the signal line is an indication of rise in price in near future. Similarly, the crossover of indicator curve from above the signal line is an indication of fall in price in near future.

Percentage price oscillator (PPO)

PPO is measured by computing the absolute difference of shorter exponential moving average and longer exponential moving average as a percentage of longer moving average.

The functioning of PPO is almost similar to MACD and it is used in order to cross check the performance of its cousin indicator MACD.

Formula:

$$\text{PPO} = \frac{\text{m period EMA} - \text{n period EMA}}{\text{n period EMA}} \times 100 \ \ (\text{m}<\text{n})$$

PPO is also a momentum oscillator that oscillates above and below the zero line. Since the shorter moving average lags lower than the longer moving average it has the much effect on PPO than longer moving average. The values of PPO would be positive when the shorter moving average is higher than the longer moving average and vice versa.

Similar to MACD, PPO also generates signals when there is center line crossover, signal line crossover, and divergences. Center line crossover occurs when the value of PPO is zero that is when both the shorter moving average and longer moving averages are equal. When the plot of shorter moving average curve crosses the longer moving average from above then it is an indication of fall in price in near future. Similarly, when the shorter moving average curve crosses the longer moving average curve from below then it is an indication of rise in price in near future. Thus the indicator generates a signal to bid a share when it is positive and offer a share when it is negative.

The default settings used for the calculation of PPO are 12, 26, and 9. The 12 periods moving average, 26 periods moving average are used to measure the values of percentage price oscillator and the 9 periods moving average is used as a signal line (stockcharts.com, 2011). The plot of signal line used along with PPO indicator is also used to generate signals of either to bid a share or offer a share. When the plot of PPO curve crosses the signal line from above then it is an indication of fall in price in near future. Similarly, when the plot of PPO curve crosses the signal line from below then it is an indication of rise in price in near future.

Similar to MACD, PPO is not suitable for the identification of overbought and oversold levels as the range of PPO is unbounded. The range of PPO indicator does not depend on the values of price unlike MACD. The range of PPO would be much less that of MACD because PPO is the measurement of MACD relative to longer moving average where as MACD is just the absolute difference of moving averages. This is a significant difference noticed in PPO when compared to MACD and it is used along with MACD in order to cross check its performance. If both generate signals to bid or offer a share then it is an indication that it is a correct prediction otherwise if one generates a bid signal and other indicator generates an offer signal then it an indication that one of them has generated a pseudo signal.

Chaikin money flow (CMF)

CMF indicator is developed by Marc Chaikin. It measures the amount of money flow volume over a specific period in general 20 or 21 days (stockcharts.com, 2011). The money flow volume forms the basis for accumulation - distribution line where the accumulation is high when the amount of money flow into the stock is high and the distribution is high when the amount of money flow out of the stock is high. Accumulation and distribution line measures the cumulative flow of money into the stock and out of a stock where as the Chaikin money flow simply adds the total money flow volume for the specific look back period typically 20 or 21 days.

This indicator is famous for the end of day trading that is the reason for the typical period used for the calculation of CMF is 20 or 21 days. Most of the indicators derived for the end of day trading only depend on closing pieces where as this indicator is derived using the maximum of available technical information that is from the closing price, lowest price, highest price and the volume of shares traded during a period.

Formula:

The following formulas adopted from www.stockcharts.com website are used for the measurement of Chaikin money flow. The values of CMF indicator oscillate above and below the zero line with the fluctuations from -1 to +1. The positive value is an indication of accumulation and a strong buying pressure, similarly the negative value is an indication of distribution and a strong selling pressure.

$$\text{money flow multiplier} = \frac{\text{(close-low)-(high-close)}}{\text{(high-low)}}$$

$$\text{money flow volume} = \text{money flow multiplier} \times \text{volume for the period}$$

$$\text{20 period CMF} = \frac{\text{20 period sum of money flow volume}}{\text{20 period sum of volume}}$$

CMF generates a signal to either a sell or buy whenever it crosses the central line or zero line, if the plot of CMF crosses the central or reference line from the below then it generates buy signal.

Similarly, if the plot of CMF crosses the reference line from the above then it generates a sell signal.

Chaikin money flow could be used along with the indicators such as RSI and William's percentage R that identifies the overbought and oversold levels of the stock. If the value of RSI is moving above 50 and CMF is positive then it is an indication of correct prediction by both the indicators. Similarly, if the value of RSI is moving less than 35 and CMF is negative then it is an indication of correct prediction by both the indicators. This is how Chaikin money flow is used along with other indicators such as RSI and William's %R to crosscheck its performance, if both generate same signals the it is a correct prediction otherwise one of the indicator should have generated a pseudo signal but it is difficult to identify which one of them has generated a pseudo signal or real signal.

Rate of change (ROC)

ROC is a pure momentum oscillating indicator that measures the percentage change in price from current price to the price n-periods ago. The plot of ROC indicator oscillates above and below the zero as a reference line. The price value rises when the indicator is positive and drops when the indicator is negative. So the crossover of the indicator from below the reference or central line is an indication of rise in price and when the indicator crosses the central line from above then it is an indication of drop in price.

The indicator generates a signal to sell the stock when its value is negative and vice versa.

Formula:

$$\text{ROC} = \left(\frac{\text{closing price of current period}}{\text{closing price n periods ago}}\text{-1}\right)\times 100$$

Rate of change indicator is not only used for the identification of the change in trend but also for the identification of overbought and oversold signals of the stock. The indicator does not have either an upper and lower boundaries but the percentage rise or drop in price is used to estimate the overbought and oversold levels.

The default parameter setting used to calculate the rate of change is 12 periods (stockcharts.com, 2011). That is the current price is compared with the value of price n-periods ago. The rate of change would be positive if the current price is greater than the price n-periods ago and negative otherwise. The positive value suggests that bull domination is going on and it is a good sign to invest on the stocks. Similarly, the negative value of indicator suggests that bear domination is going on which is a good sign to sell the stock to make profits.

Commodity channel index (CMI)

"Developed by Donald Lambert and featured in Commodities magazine in 1980, the Commodity Channel Index (CCI) is a versatile indicator that can be used to identify a new trend or warn of extreme conditions" (stockcharts.com, 2011).

The indicators is developed for the identification of cyclical turns followed by the commodities but later the technical analysts were successful in applying it to indexes and stocks.

Formula:

The following formulas from www.stockcharts.com website are used in the calculation of this indicator with the default parameter setting taken as 20 periods.

$$\text{CCI} = (\text{TP-20 period SMA of TP}) \div (0.15 \times \text{Mean Deviation})$$

$$\text{Typical Price, TP} = (\text{High} + \text{Low} + \text{Close}) \div 3$$

This indicator is derived from the technical data of highest price, lowest price, and closing price during a specific period. The average of these three prices called typical price is used for the calculation of commodity channel index value which is a ratio of absolute difference between the current typical price and the 20-periods simple moving average to the mean deviation of prices.

In order to measure the mean deviation, first subtract the most recent 20-period average of the typical price from the typical price of each period. Second, take an absolute of these values. Third, add all these values. Fourth, divide it with default parameter setting that is 20-periods. (stockcharts.com, 2011)

CCI is a momentum oscillator that oscillates above and below the zero or reference line. It is an unbounded indicator but could still useful in indentifying the overbought and oversold levels. The value of an indicator is high when its price is far above the average price and it is low when the price is far below its average price. The constant is set to 0.15 such that about 70 to 80 percent of the indicator values would fall in the range of -100 to +100 (stockcharts.com, 2011). The value of indicator below -100 is an indication of downtrend and above 100 is an indication of upward trend in price. The value of CCI above +100 is useful in the identification of overbought and value below -100 is useful in the identification of oversold levels.

Relative strength index (RSI)

J. Welles Wilder developed the RSI indicator which is a momentum oscillator that measures the speed and the change in price movements (stockcharts.com, 2011). The important thing to know about this indicator is that the value of RSI above 70 is an indication of overbought level and the values of RSI below 30 is an indication of oversold level.

Therefore a sell signal is generated if the value of indicator is less than 30 and a buy signal is generated if the value is greater than 70. It is a reliable indicator but does not inform the trader about how long would the stock be in overbought or oversold level. It infers that a stock could be in an oversold or overbought period for long periods.

For instance, if the stock X has an RSI value of 70 in August, 2011 which is an indication of overbought and suggests the trader to sell his stock. The stock X might have the RSI value above 70 until November, 2011 but the RSI indicator fails in identification of the change in overbought and oversold levels.

That is a reason RSI is used in conjunction with other indicators such as William's %R that forecasts the overbought and oversold levels to crosscheck its performance. If both the indicators generate a similar signal of either to bid or offer a share then the forecast is real otherwise if one indicator generates a signal to bid and other indicator generates a signal to sell then it is an indication of pseudo signal generation.

Formulae:

The following formulas are used from www.stockcharts.com website to estimate the values of relative strength index that ranges from 0 to 100.

$$\text{RSI} = 100 - \frac{100}{1+\text{RS}}$$

$$\text{RS} = \text{average gain} \div \text{average loss}$$

$$\text{first average gain} = \text{average of gains over past n periods}$$

$$\text{first average loss} = \text{average of losses over past n periods}$$

$$\text{average gain} = \frac{(\text{previous average gain} \times (\text{n-1}) + \text{current gain})}{\text{n}}$$

$$\text{average loss} = \frac{(\text{previous average loss} \times (\text{n-1}) + \text{current loss})}{\text{n}}$$

The default setting for n is taken as 14.

Williams percent R (WPR)

WPR is a momentum indicator that reflects the level of closing price relative to the highest high for the specific look-back period. It oscillates in the range of 0 to -100. The indicator readings from 0 to -20 are considered overbought and readings from -80 to -100 are considered oversold.

WPR is a measurement of closing price relative to highest high for the specific look back period (stockcharts.com, 2011). It oscillates in the range of 0 to -100 where the readings less than -80 are considered as oversold and the readings above -20 are considered as overbought.

WPR generates a signal to bid when the value is above -20 and generates a signal to offer when the value is less than -80. Its functioning is similar to that of RSI indicator and it is used to crosscheck the performance with RSI indicator.

WPR compares current price to the recent range indicating whether bulls could close the market at a higher price level or bear could close the market at a lower price level. If bulls cannot close the market at a higher level then bulls are weak and results in fall in price. Similarly, if the bears cannot close the market at a lower level then bears are weak and results in a rise in price.

Formula:

The following formula from www.stockcharts.com website is used to measure the value of WPR indicator which is also used to estimate the overbought and oversold levels.

$$\text{WPR} = \frac{(\text{highest high - close})}{(\text{highest high - lowest low})} \times \text{-100}$$

The default parameter setting is chosen as 14-periods (stockcharts.com, 2011).

A 14-period WPR is measured on taking a ratio of the absolute difference of highest value over last 14-period and closing price to the absolute difference of highest high and lowest low values over the last 14-perriods.

Highest high is the highest high over the specific look back period.

Lowest low is the lowest low over the specific look back period (14).

1.6 Traditional Time Series

The traditional time series approach analyzes historical data and predicts the orientation of future price as a linear function of historic data. In econometrics, an application of mathematics and statistics to the study of economic and financial data, there are two basic types of time series prediction strategies simple regression and multivariate regression (Kalyvas, 2001).

Regression analysis is used to forecast a dependent variable using other independent variables. Simple or linear regression is an approach that use just a single independent variable whereas multivariate regression model is a technique that use more than one (>1) independent variables to derive the dependent variable.

1.7 Machine Learning Algorithms

Several machine learning algorithms such as boosting, decision stumps, support vector machines, k-nearest neighbors, genetic algorithms, reinforcement learning, neural networks and so forth are applied for stock trading to predict the fluctuation of stock prices before the actual event of an increase or decrease in the price occurs (Shah, 2007) (Kalyvas, 2001).

1.8 Correlation of Stocks

Correlation determines the degree to which the two stocks are associated which is computed using the correlation coefficient that ranges from -1 to +1. Correlation coefficient of stocks X and Y is the ratio of covariance of X and Y to the product of stand deviations of X and Y.

$$\text{Correlation}\ (X,Y) = \text{Covariance}\ (X,Y)\ /\ SD(X)\ SD(Y)\ n$$

$$\text{Covariance}(X, Y) = \sum(X(i)\text{-}XMA)\ (Y(i)\text{-}YMA)$$

$$SD(X) = \sqrt{\frac{\sum_{1..n}(X(i)\text{-}XMA)^2}{n\text{-}1}}$$

Where X and Y are the closing prices of stock X and Y. SD(X) and SD(Y) are the standard deviations for the stocks X and Y. n is the total number of periods, which would be 1 for end of day trading with the time frame as a single day and 7 for the end of day trading with the time frame taken as one week. XMA and YMA are the moving average estimates for the stocks X and Y using their closing price.

The extreme values +1 and -1 of correlation would represent the strong association among two stocks, +1 indicates the two stocks are moving in the same direction while -1 indicates the two stocks move in the opposite direction. The value 0 is an indication that both stocks are said to have no correlation or random.

We have compared the performance of a single stock on preparing the data set with the nine chosen technical indicators and also on preparing a data set with the correlation of a stock with the number of features equal to the product of nine and the number of related stocks, m.

We first acquired m related stocks that exhibit high correlation factor with the forecasting stock. One of these stocks will always be the target stock itself as it will have perfect correlation with itself (Choudhry & Garg, 2008). Our stock trading model using XCS takes a series of nine technical indicators for each correlated stock as candidates for input features. The set of nine times m features combined together could be used as the candidate features for a stock. The XCS model for stock trading is tested using both the modes and found that XCS could generate higher returns with the correlation features instead of a single stock features.

The XCS trading model performed better with the correlation features because the variation of prices in one correlated stock could be propagated to other correlated stocks due to emotional or psychological factors. The current data set with nine times m candidate features would reflect the changes in future price of the stock using information from the correlated stocks. For example, assume that X and Y are two small scale software companies and the positive statement from a finance minister of a nation related to that sector would indeed cause the prices to rise in future but one might rise in advance and other a bit later in time period.

Correlation mechanism is used as a sort of advanced portfolio management technique in order to earn optimum returns rather than estimating profits using a single stock features.

The correlation of a stock is also available on google finance. It supplies the top ten correlated stocks along with the percentage change in price for each of the correlated stock and could observe that all correlated stocks would either go along a bullish pattern or a bearish pattern, which is an indication that all correlated stocks pursue a similar trend.

The following is a real time instance for the correlation of IBM stock. It shows the top ten stocks or companies that are much associated to IBM.

	Company name	Change	Chg %
IBM	Intl. Business Machine	+2.20	1.33%
HPQ	Hewlett-Packard Company	+0.51	2.16%
ORCL	Oracle Corporation	+1.14	4.30%
MSFT	Microsoft Corporation	+0.49	1.92%
GOOG	Google Inc.	+11.85	2.27%
CSCO	Cisco Systems, Inc.	+0.60	3.93%
INTC	Intel Corporation	+0.54	2.76%
AAPL	Apple Inc.	+4.19	1.10%
DELL	Dell Inc.	+0.30	2.11%
EMC	EMC Corporation	+0.71	3.32%
SAP	SAP AG (ADR)	+2.01	3.95%

1^{st}, 2^{nd} column represents the symbol and name of a correlated stock

3^{rd} column represents the change in price over the recent quarter

4^{th} column represents the percentage change in price

It can be noticed from here that the correlation of IBM with IBM itself will have the maximum +1 correlation coefficient and all other correlated stocks shown are moving in the same direction as the IBM stock which means the correlation coefficient of IBM with each of these companies would be somewhere in the range of greater than 0 and less than 1.

Chapter 2 - Background and Related Work

2.1 Financial Terms

The term financial instrument is defined as a tradable asset or negotiable item such as security, commodity, or index that underlies a derivative. It is a manner through which something of value is transferred, held or accomplished.

A security has some financial value and it is broadly categorized into debt securities (such as banknotes, bonds and debentures) and equity securities (such as common stocks and derivative contracts). Commodity is a raw material that is bought and sold commercially in large quantities.

A stock also known as equity is a kind of security that expresses ownership in a corporation and represents a claim on part of the assets and earnings of a corporation. Common stock and Preferred stock are the two main divisions of stock.

A common stock is a specific kind of security that represents the ownership in a corporation. It usually entitles the owner to vote at shareholders meetings and to receive dividends. People that hold common stocks exert control through electing a board of directors and voting on corporate policy. These shareholders are on bottom of the priority ladder for ownership structure. In the event of liquidation these shareholders have rights on the assets of a corporation only after bond holders, preferred shareholders and other debt holders have been paid in full.

A preferred stock is a class of ownership in a corporation that has a higher claim on the assets and earnings than common stock. It generally does not have voting rights, but has a higher claim on assets and earnings than the common shares. For example, owners of preferred stock receive dividends before common shareholders and have priority in the event that a company goes bankrupt and is liquidated. The precise details as to the structure of preferred stock are specific to each corporation. However, the best way to think of preferred stock is as a financial instrument that has characteristics of both debt (fixed dividends) and equity (potential appreciation).

Equity is a stock that represents an ownership interest. It is the amount of the funds contributed by owners (stockholders) plus retained earnings (or losses) on the balance sheet of a firm.

A market also known as stock market or equity market is defined as a stock exchange or over-the-counter market that issues and trades stock. It is one of the most critical areas of economic system as it provides companies with access to capital and investors with a slice of ownership in the corporation and the potential of gains based on the future performance of corporation.

Exchange is an entity that provides service for stock brokers and traders to trade stocks, bonds and other securities. The common securities traded on a stock exchange include shares issued by companies, unit trusts, derivatives, pooled investment products and bonds.

Portfolio is an investment and risk limiting scheme where the chance of risk can be reduced on owning shares of multiple assets. The assets could include bank accounts, stocks, bonds, options, warrants, gold certificates, real estate, future contracts, production facilities or any other item that is expected to return its value.

Open: trade price of the first transaction during a day.
Close: trade price of the last transaction during a day.
Bid Price: the price at which an investor can sell securities.
Offer Price: the price at which an investor can buy securities he/she holds.
Bull Market: a period when most stocks are increasing in value.
Bear Market: a period when most stocks are declining in value.
Index: a composite representing the value of a group of stocks.
Liquidity: a measure of the number of shares, or dollar value of shares traded.
Volume: number of shares traded during a specified time, usually one day.
Portfolio: a group of stocks, mutual funds, or other securities.

2.2 Random Walk Process

The objective of stock traders is to earn profits through the prediction of future price movement using various mechanisms available but according to random walk theory the price movements are unpredictable. The random walk experiment of Maurice Kendall, observed in 1953, states that price and index movements are absolute random which means future price and index values are independent of historical information that makes stock market efficient (Kendall, 1953). The stock markets could be efficient even with the irrational investors and also with the stock prices exhibiting greater volatility than could apparently be explained (Malkiel, 2003). Some financial analysts believe that movement of price is controlled through a random walk process assuming there is no correlation between past price, volume, and other indicators thus it is regarded as unpredictable. The unpredictable behavior of price appears to establish the irrational behavior of stock market however on deep examination it became evident that random walk process indicates a well operating and rational one (Bodie, Kane, & Marcus, 2008).

2.3 Efficient Market Hypothesis (EMH)

EMH proposes that current stock price reflects the entire available information about the firm and market thus it is impossible to make profits using this information. EMH has three different versions based on the kind of information available which are weak level hypothesis, semi strong hypothesis and strong level hypothesis (Clarke, Jandik, & Mandelker, 1999).

Weak level hypothesis suggests that current stock price reflects the entire available trading data such as historical price, and volume. It is impossible to beat or outperform stock market through technical analysis when the past data of price and volume is available for entire public then there is no use of processing and analyzing this data to predict the future stock price because whenever a trader predicts the future stock price using some combination of technical indicators other traders could also use a similar kind of approach to predict the future stock price. It suggests that technical analysis is useless (Clarke, Jandik, & Mandelker, 1999).

Semi strong hypothesis suggests that the current stock price reflects the entire data available to public. The public available information signifies both technical data (such as price, volume) and fundamental data (such as assets, earnings, dividends, income, revenue, patents held, product line, accounting practices, etc. which are derived from balance sheets and income statements of a firm). It is based on the assertion that one should not be able to make profits using something which is known to public (Clarke, Jandik, & Mandelker, 1999). This principle suggests that both technical analysis and fundamental analysis as useless.

Strong level hypothesis suggests that stock prices reflect all information available to both public and private reflecting both weak level hypothesis and semi-strong level hypothesis. The most important difference between semi-strong level hypothesis and strong level hypothesis is that in the latter case no trader should be able to make profits even while trading when the information in not known to public but just known to the management itself. Given this assumption none of the investor would be able to make profits above the average investor even if he was given both public and private (also known as inside information) information that would not give scope to generate profits (Clarke, Jandik, & Mandelker, 1999).

In concise, it can be established that when there is some exploitable information then the mass of intelligent investors would earn profits from such prediction which move stock price and cause the trading strategy to self destruct (Chauhan, 2008).

2.3.1 Common Misconceptions

Myth 1: EMH asserts that stock traders cannot beat or outperform the market. However the investors like Warren Buffett and George Soros were able to forecast the stock market and make huge profits (Clarke, Jandik, & Mandelker, 1999). Therefore, EMH must be incorrect.

Myth 2: “EMH claims that financial analysis is pointless and investors who attempt to research security prices are wasting their time” (Clarke, Jandik, & Mandelker, 1999). Therefore, EMH must be incorrect.

Myth 3: EMH claims that even the recent available information is completely incorporated in current stock price but the fluctuations in price could be noticed every day, hour, minute, and second even without much difference in the available information (Clarke, Jandik, & Mandelker, 1999). Thus, EMH must be incorrect.

Myth 4: EMH assumes that all stock traders to be skilled, informed, and are capable of analyzing the constant flow of new information however there were a number of common people who does not have sufficient knowledge on prediction methodologies but does trading (Clarke, Jandik, & Mandelker, 1999). Therefore, EMH must be incorrect.

2.4 Related Work

E. Fama's efficient market hypothesis (Fama, 1970) and Martingale Model (Samuelson, 1965) (Mandelbrot, 1966) rules out the application of both technical analysis and fundamental analysis prediction strategies to the stock market as the technical data such as previous price and volume information, fundamental data such as income statement, dividends, assets, etc. available to public is incorporated in current price and thus become useless for the prediction of future price. Even the private or inside information which is just known to management or insiders of firm but not known to general public is not useful for the prediction of future price according to the efficient market hypothesis (EMH).

There are proponents on both sides where some financial analysts believe the stock market is predictable using the fundamental data and technical data whereas others consider the market as unpredictable and the price or index follows a random walk process. Burton J. Malkiel in his random walk experiment (Burton, 1999) proposed that stock prices are absolute random which means the price on next day is completely independent of current and previous price information. Andrew Lo and MacKinlay (Andrew Lo & Mackinlay, 2002) established that prices or indices are predictable which follows a long-term trend. Andrew Lo and MacKinlay took the 23 years of US stock data from the period of 1962 to 1985 for their research with the time frame as one week. Their investigation proved that stock prices are predictable with the identification of long-term trend and ruled out the random walk process (Chauhan, 2008).

The work done by An Pin Chen and Mu Yen Chen depicted that stock price or index fluctuations could be predicted using the extended classifier system and the technical indicators which would generate good profits with the accurate predictions (Chen & Chen, 2005) (Chauhan, 2008). Schulenburg developed a learning classifier system (LCS) model in her PhD research and tested its performance on several combinations of technical indicators (S & P, 2001). Stone worked on the foreign market exchange prediction with ZCS algorithm in his PhD (Stone & Bull, 2004). Chen Lin used XCS to predict the fluctuations of future stock price and index. This model used moving averages of price and volume information to construct the message (Chen, Lin, & Chen, 2007). Gershoff, M. and S. Schulenburg described the collective behavior of extended classifier system to achieve higher accuracy in prediction with an objective to make optimum returns (Gershoff & Schulenburg, 2007) (Chauhan, 2008).

2.5 Learning Classifier Systems

A learning classifier system or LCS is a machine learning technique that combines reinforcement learning, genetic algorithms, and heuristics to develop an expert system. LCS is adaptive in the sense that it learns from experience through receiving a feedback from environment or simulator for each input in the form of payoff. The payoff is a numerical value that inform LCS agent whether the prediction or classification proposed is right or wrong, the negative value would represent the prediction is false and true otherwise.

LCS was first invented in 1975 (Sigaud & Wilson, 2007) but several variants of LCS were proposed thereafter based on the fitness measurement, application of GA, and learning algorithm, etc. In general all LCSs contain a set of condition-action rules called classifiers that represents the entire knowledge domain. The rules are updated using payoff received from environment also the fittest rules are evolved and added to the existing knowledge domain through replacement with the rules that have least fitness values in order to maintain a fixed number of total rules or classifiers. This set of rules that represent a knowledge domain is known as population set [P].

There are several models of LCSs as a result multiple definitions exist for the learning classifier system (Holmes, Lanzi, Stolzmann, & Wilson, 2002). Nevertheless all LCS models more or less comprise the four main components the population set, performance component, reinforcement component, and rule discovery component (Holmes, Lanzi, Stolzmann, & Wilson, 2002).

Performance component controls the interaction of classifier system with environment (Holmes, Lanzi, Stolzmann, & Wilson, 2002). It is like an intelligent agent though it is less dependent on the problem domain and it is a rule (or classifier) based message-passing, highly parallel, and highly standardized expert system (Booker, Goldberg, & Holland, 1989).

Reinforcement or credit assignment component distribute the actual payoff to the rules that are accountable in choosing the best action to execute (Holmes, Lanzi, Stolzmann, & Wilson, 2002). [P] has rules with different fitness value estimates, the one with high fitness value indicate a useful rule and the one with least fitness value could be incorrect rule. The Bucket-Brigade or Q-Learning Reinforcement algorithms are used to update the rule fitness and other parameters of a rule using the actual payoff received from the environment or simulator.

Rule discovery component identifies best fittest rules, create clones of those and apply genetic operations (such as mutation and crossover) to evolve the child classifiers. These child classifiers would represent fitter or useful rules whose fitness is the average of the parent classifiers. The goal of applying genetic algorithm is to evolve best classifiers in the rule set [P].

Each rule or classifier has two parameters which are the prediction (also called strength) and fitness (Holmes, Lanzi, Stolzmann, & Wilson, 2002). Holland's LCS has just a single parameter for both the fitness and strength measures but the later models use both. The estimate of fitness indicates the usefulness of a rule while the estimate of prediction indicates the expected payoff that the agent would receive on selecting that particular classifier. Fitness determines the quality of knowledge about the problem domain that a rule express and it is the basis for the evolution of rules using genetic algorithms. A rule with higher fitness suggests that it has a better quality of information and a rule with lower fitness suggests that it has no or a little useful information (Holmes, Lanzi, Stolzmann, & Wilson, 2002).

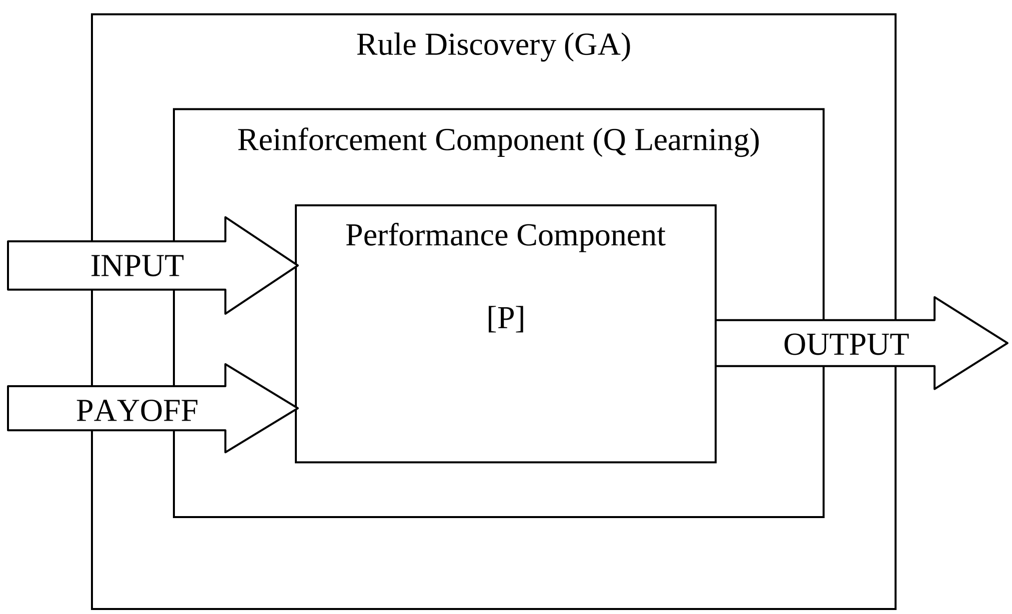

Figure 4 - General LCS Framework

2.5.1 Pittsburgh vs. Michigan styles of LCS

Learning classifier systems is structured as the so called Michigan style and Pittsburgh style approaches (Sigaud & Wilson, 2007).

Smith from the University of Pittsburgh proposed a LCS (Smith, 1980) that adapts only through the application of genetic algorithms on the rule set [P]. The population set has multiple individuals where each individual is a classifier that has multiple classifiers and each classifier has multiple rules. Each classifier is capable of proposing a best possible action to execute.

Holland and his PhD students at the Michigan State University proposed a LCS that combined the genetic algorithm and reinforcement learning to predict the best possible action to execute (Sigaud & Wilson, 2007). The population set has multiple individuals where each individual has a single classifier and each classifier represents a single condition-action rule. The population set as a whole is used for the prediction of best action to execute in order to receive optimum returns or payoff.

Pittsburgh style approach is getting engaged into the wider evolutionary computation domain as it is based on the traditional genetic algorithms. GABIL is the example for this kind of approach.

Michigan style approach became the standard LCS which applies the reinforcement learning along with the genetic algorithms to adapt to the fittest rules in the population set. ZCS and XCS are the popular examples of this kind of approach.

ZCS is based on payoff (strength) while XCS is based on the accuracy of payoff predictions. ZCS fitness is directly proportional to the payoff while XCS fitness measure is directly proportional to the accuracy or inversely proportional to the prediction error.

When compared to ZCS, XCS does perform much better because it computes the payoff using accuracy of payoff predictions instead of payoff itself and also due to the application of genetic algorithm on only the active rule set [A] instead of applying it on the entire population set [P].

Our trading model is developed using the Wilsons XCS algorithm (Martin & Stewart, 2000) and applied to the stock market simulator which is represented using the technical indicators.

2.6 Genetic Algorithms (GA)

A genetic algorithm belongs to the larger class of evolution algorithms that often perform well in estimating the solutions to all kinds of problem domains and it is used to evolve the accurate candidate solutions for the optimum returns using selection, inheritance, mutation, and crossover. GA is often used in optimization and search solutions using techniques inspired from the natural evolution process.

GA depends on four analogies with their biological counterparts that use [1] a code used to represent the problem domain, the *chromosome* or *genome* or *genotype*, [2] the simple transformations applied on that code, *genetic operations*, [3] the expression of a solution from the code, the *genotype-to-phenotype* mapping, and [4] a solution selection process, the *survival of fittest* (Sigaud & Wilson, 2007).

GA is applied on a population of strings (conditions or chromosomes) that encode candidate solutions (actions or phenotypes) to evolve towards better rules. In general, the individuals are represented as a binary string of 0s and 1s.

In general, Genetic algorithms are applied on a population [P] of classifiers or rules to evolve better rules in order to make optimal returns. Each classifier is a condition-action rule that consists of four basic parameters condition, action, prediction, and fitness. In biology, the term condition is used to represent the current input situation of the environment is analogous to chromosome which is usually expressed as a binary string of 0s and 1s. Similarly, the term action which is analogous to phenotype in biology is an indication of proposed action for the perceived input from the environment. Prediction represents the strength or payoff expected when the classifier or rule is employed. Fitness estimate indicates the usefulness of a rule or classifier.

Example:

Chromosome or Condition: 101011001

Phenotype or Action: 10

Classifier or Rule: 101011001: 10   16.2 0.09

The above instance of a classifier asserts that for a given input condition 101011001 and if the chosen action for execution is 10 then the expected payoff is 16.2 units and the fitness measure is 0.09 units.

A single rule or classifier could be used to hold information attained from large number of inputs perceived from the environment or simulator. So, In order to satisfy this criterion the classifier condition is represented as a string of ternary alphabets of 0s, 1s, and #s.

Where initially the each bit which is analogous to a gene in biology which indicates whether gene does exist in a genome or not. Similarly, Here it indicates whether the candidate feature used to represent a simulator or environment is present or not. But in the modified version of a condition string an extra character # is added to include a do not care along with 0 and 1 that suggest whether the feature or gene is present or not.

Example:

Classifier or Rule: 101011###: 10   16.2 0.09

The above instance of a rule would influence the action of execution for the following eight input situations perceived from the environment.

101011000
101011001
101011010
101011011
101011100
101011101
101011110
101011111

A classifier with this condition 101011### is selected for all the above input situations because of the do not care symbol embedded in the rule. This assists the rule 101011### to adapt from multiple input strings and gets updated using the payoff received for the proposed action to execute. Since there were only two possible actions for each candidate feature the total number of inputs that a rule with n # symbols would act upon are 2^n.

The action string is also represented using 0s and 1s. The number of bits is chosen depending on the possible number of actions for execution. If m is the size of an action string then 2^m would indicate the maximum number of actions that are possible for execution.

Fitness function is determined so as to get optimum returns using the greedy approach. It is dependent on the problem domain. Fitness measure is the basis for selection mechanism, a rule with least fitness value is selected for deletion and a rule with high fitness value is selected for survival using a roulette wheel selection mechanism.

The two most popular genetic operations used for the reproduction are mutation and crossover. Single point crossover is used in all the traditional LCSs which divides a chromosome into two parts at a random location of the condition string and generates a new chromosome through inverting the subparts from different parents. Similarly, a multi way crossover could be derived from the multiple partitions. Mutation is also a genetic operation that changes the existence of genes in a genome. Higher the probability of mutation then higher would be the change in number of bits in the chromosome or condition of a rule and vice versa.

2.7 Markov Decision Process and Reinforcement Learning

2.7.1 Markov Decision Process (MDP)

A markov decision process consists of the following components (Sigaud & Wilson, 2007)

a) A finite set S of discrete states s of an agent
b) A finite set A of discrete actions a
c) A transition function Pf: $S\ X\ A \rightarrow \pi(S)$ where $\pi(S)$ is the set of probability distributions over S. It is also referred as a policy function.
d) A reward function Rf: S X A $\rightarrow R$

The probability distribution of a transition function Pr ($s_{t+1}|s_t, a_t$) suggests that agent reached to next state s_{t+1} from state s_t on executing the action a_t.

The reward function indicates the payoff received from environment when the agent in state s_t executes an action a_t. It is a mapping of (s_t, a_t) tuple to the payoff that represents when the agent in state s_t and it chooses an action a_t for execution then R is the reward or payoff received.

According to the markov decision property the next state s_{t+1} just depends on the previous state and action chosen (s_t, a_t) but does not depend on other historical states and actions.

Thus, $Pr(s_{t+1}|s_t, a_t) = Pr(s_{t+1}|s_t, a_t, s_{t-1}, a_{t-1}, s_{t-2}, a_{t-2},$ and so on up to $s_0, a_0)$

The aim of MDPs is to maximize the cumulative sum of accumulated rewards. The discount factor $\gamma\epsilon[0,1]$ is used to estimate the future rewards taken into account in the computation of accumulated rewards at time t as follows (Sigaud & Wilson, 2007):

$$Rc_{\pi}(t)=\sum_{k=t}^{T_{max}} \gamma^{k-t}\, r_{\pi}(k)$$

Where $r_{\pi}(k)$ indicates an immediate payoff received at time k when the agent follows a policy π and T_{max} could be either finite or infinite.

MDPs can be solved using dynamic programming methods which introduce a value function v^{π} where $v^{\pi}(s)$ acts for each state the payoff that an agent could anticipate when it follows the policy π (Sigaud & Wilson, 2007).

Instead of the value function v^{π}, it is much useful in defining an action-value function Q^{π} where $Q^{\pi}(s,a)$ represents the accumulated rewards that the agent could anticipate when it performs an action a in state s following a policy π (Sigaud & Wilson, 2007).

The direct relation between value function and action-value function is established as follows $v^{\pi}(s)=\max_a Q^{\pi}(s,a)$ and the mathematical equation of value function is defined as $\forall\, s \in S, V^{\pi}(s)= \sum_a \pi(s_t,a_t)[R(s_t,a_t)+\gamma \sum_{s_{t+1}} P(s_{t+1}|s_t,a_t)V^{\pi}(s_{t+1})$ (Sigaud & Wilson, 2007).

2.7.2 Reinforcement Learning (RL)

RL maps the current input situation to the output or action so as to receive the maximum reward from the environment. Unlike supervised learning, where the agent is given training information that contains small percentage of instances that map form input to output, the reinforcement learning is not provided with such information rather it learns from the feedback received from the environment in the form of a numerical value or reward for each input.

RL is often used for the interactive problems, the agent is not given information about the action to execute rather indentifies which action result in highest reward and learns through trial and error. In most interesting and challenging cases the action chosen would not affect just the immediate reward but also the subsequent future outcome and reward (Sutton & Barto, 1998). The trial and error search and delayed reward are the important characteristics of reinforcement learning (Sutton & Barto, 1998).

RL captures the agent interactions with the environment with a goal to achieve optimal rewards. The agent should be able to perceive the state of environment or current input situation and choose the best plausible outcome to execute. The agent learns through experience, so it does not have sufficient knowledge at the initial phase as a result the predictions might be wrong but the agent improves its features through receiving feedback from an environment in the form of a numerical value or reward or payoff. The learning parameters and features were updated using the payoff received and thus agent would give a correct classification or prediction in future for the same input situation even though the initial prediction was wrong. The value of payoff determines whether the prediction is correct or wrong, in general -1 (negative) and +1 (positive) are used for the correct and wrong estimates.

Therefore an important challenge of the agent is to maintain a tradeoff between exploitation and exploration. To receive optimal reward the agent must prefer actions that it has tried in the past and found to be effective in producing reward (Sutton & Barto, 1998). The agent should explore all different input situations that are not perceived so far and also exploit the existing knowledge domain through updating the learning parameters using payoff received from environment.

RL has three components that are i) environment is specific to a problem domain that an agent interacts with and it is represented using the features, ii) agent is a learner or decision maker, and iii) action is a plausible outcome that an agent suggests for execution. The agent has four basic components which are a policy, a reward function, a value function, and a model of the environment (Sutton & Barto, 1998).

The policy is a decision making function or lookup table or extensive computation of agent that decides an action to execute at all different situations that it encounters (Sutton & Barto, 1998). In other words, it is a mapping from states of environment to actions to be chosen when the agent is in those states (Sutton & Barto, 1998). It is considered as the core of reinforcement learning approach as it alone is sufficient to determine the behavior and the other components serve just to improve the policy (Sutton & Barto, 1998).

A reward function determines the goal in a reinforcement learning problem (Sutton & Barto, 1998). In other words, it maps each state - action pair with a numerical value called reward that indicates the intrinsic desirability of that state - action pair in future (Sutton & Barto, 1998). The objective of the reward function is to maximize the sum of cumulative rewards in long run. It identifies what events are good and bad using the sign of a numerical value (Sutton & Barto, 1998). If the agent has selected an action that results in negative reward then the agent learns using this reward supplied and does not choose this action again future rather go for other actions. In general the reward functions are stochastic (Sutton & Barto, 1998).

A value function defines what is good in long run unlike a reward function that identifies what is good in an immediate sense (Sutton & Barto, 1998). In other words, it is the measure of a future rewards accumulated from the current situation or state (Sutton & Barto, 1998). For example, a state or situation could have a small immediate reward but that does not mean it also has a lower value. The value would be higher when the future rewards are high and even with the lower immediate reward. To make analogies with human the high rewards are like pleasure and low rewards are like pain (Sutton & Barto, 1998).

Rewards are the numerical values provided for each executed outcome at the current state or situation of environment. Without rewards there could not be values which are the prediction of rewards and the sole purpose to compute values is to maximize accumulated future rewards (Sutton & Barto, 1998).

RL has two main strategies, one thread trial and error search runs through some of the earliest work in artificial intelligence that led to the revival in 1980s (Sutton & Barto, 1998). The second thread deals with the problem of optimal control and its solution using value functions and dynamic programming because it is considered as the sole feasible approach for solving general stochastic optimal control problems (Sutton & Barto, 1998).

All RL methods are structured around estimating the value functions but it is not only the single strategy to solve these problems. There were other search methods such as genetic algorithms, genetic programming, stimulated annealing, and other function optimization methods that could be used to solve these problems (Sutton & Barto, 1998). These methods search in the space of policies without even computing the value functions (Sutton & Barto, 1998). These methods are known as evolutionary algorithms because their operation is analogous to the manner that biological evolution generates child organisms through reproduction from the parent organisms. These algorithms have an advantage when the agent could not accurately perceive the state or situation of the environment (Sutton & Barto, 1998). However, RL involves learning through interaction with the environment while GA or other search methods using evolution does not learn through interaction with the environment.

The reinforcement learning and evolutionary methods both share a lot of common features and thus could work together although we consider evolutionary methods are not well suitable for reinforcement problems (Sutton & Barto, 1998). The combination of reinforcement learning, genetic algorithm, and other heuristics is known as a learning classifier system that has the advantages of both RL and GA.

This thesis just reported in brief about the reinforcement learning approach. For supplement reading on different forms of reinforcement learning problem and MDP refer to the book of Richard S. Sutton and Andrew G. Barto on Reinforcement Learning: An Introduction (Button & Barto, 1998).

Chapter 3 - Experimental Setup

3.1 Software Requirements

3.1.1 Java

A complete object oriented programming language, java, is used for the implementation of Wilson's XCS algorithm. The entire algorithm of learning classifier system is divided into seven classes for the representation of an environment (stock market simulator), input, classifier (rule), classifier set (rule set), constants (default classifier and learning parameter settings), system prediction for computing the best action winner, and XCS.

3.1.2 Eclipse IDE

Eclipse is used as an interface development environment for the java programming language that assisted in writing the code with minimum effort, detecting the errors and correcting them.

3.1.3 MS Excel (mathematical formulas and if-then-else constructs)

Microsoft Excel is used for the preparation of data set. The mathematical formulas feature available in excel sheet is used to measure indicators using the past technical data (open price, close price, low price, high price, and volume of shares) regarding a firm is downloaded from the google finance in to an excel worksheet. Indicators are computed using the formulas mentioned in first chapter. The mathematical formula used to measure an indicator can be extended to other locations in the same column just through dragging from initial location till the end. The signals to bid and offer are generated using the if-then-else construct provided in the excel worksheet.

3.2 Dataset Formulation

The historical quotes of a stock or firm are downloaded from google finance and then processed to compute technical indicators using mathematical formulae as described in the first chapter. The estimate of each indicator would guide the investor to either bid a stock or offer a stock in order to generate profits. The signal bits generated through technical analysis are combined together to form a bit string of fixed length equal to the number of indicators used for the analysis. The bit string represents the current situation of the stock that is given as input to the learning classifier system model which selects a best possible action for execution.

The dataset downloaded from the google finance website to an excel worksheet would contain just the basic technical information about a stock or the firm. The technical information for the EOD trading would include open price, lowest price, highest price, closing price and volume of shares traded. In the computation of various technical indicators the closing price is taken into consideration for most of the indicators in EOD trading estimation.

The three arguments that need to be supplied in the extraction of technical data from google finance are time frame, start date and end date. The time frame is a difference in time gap between successive data points where as start date and end date data indicate the total number of data points to extract that represents a time period during which the technical data should be considered for processing and analyzing. The time difference could be in minutes, hours, days, weeks, or months depending on the investment plans. In general, if the time gap among successive trading periods is selected in minutes then technical analysis could be done for intraday trading and if the time frame chosen is in days then it could be applied to EOD trading. Upon considering moving average of a closing price rather than closing price the technical analysis could be applied for trading long-term investments.

Our XCS model developed for stock trading is tested using the EOD technical data where the indicators use price and volume information that is processed to guide the investors in selecting a best action to execute in order to make optimum returns.

There are several tools available to estimate the values of these technical indicators each agent use their individual approach to generate the bid and offer signals. We have not used these tools to generate a real time dataset (or test bed) rather measured the values of these indicators using mathematical formulas as explained in the first chapter. The crucial idea behind generating the signal bits using heuristics is taken from “Automated Stock Trading and Portfolio Optimization Using XCS Trader and Technical Analysis” (Chauhan, 2008).

3.2.1 Moving Average (MA)

The moving average is used in the following two modes, one for SMA and other for EMA. If the current closing price value is greater than some x-period moving average then it is a good sign to bid a stock as it is an indication that the price value is following an uptrend and vice versa. At the same time if the shorter moving average is higher than the longer moving average then it is a good sign to bid a stock again as it an indication of rise in closing price and vice versa.

```
if (5 period SMA [stock]≥close) then signal=0 else signal=1
if (50 period EMA [stock]≥ 5 period EMA [stock]) then signal=0 else signal=1
```

The signal bit indicates a prediction of the indicator to either bid a stock or offer a stock where the bit 0 indicates a sell signal and bit 1 indicates a buy signal.

The simple moving average is a good indicator for stocks that do not change its value to a greater extent even after longer periods where as for the stocks that change price often it is the exponential moving average that work better as it gives higher weights to the recent prices.

3.2.2 Commodity Channel Index (CCI)

CCI is a band oscillator. CCI above +100 indicates overbought stock and offer signal is triggered. CCI less than -100 indicates oversold stock and bid signal is triggered. The values of CCI in between [-100,100] does not give clear signal of either to bid a stock or offer a stock. So heuristics are applied in such a scenario that is if the current CCI is greater than past ten periods moving average of CCI then a bid signal is triggered and vice versa.

```
if (CCI[stock]≤-100 {
signal=1
}
else if(CCI[stock]≥100){
signal=0
}
else if(CCI[stock]>CCIMA[stock]){
signal=1
}else{
signal=0
}
```

3.2.3 Chaikin Money Flow (CMF)

A sell signal is generated while the estimate of CMF is negative and vice versa.

```
if(CMF[stock]<0) {
signal=0
else
signal=1
}
```

3.2.4 Moving average convergence divergence (MACD)

MACD indicates bullish while positive and bearish while it is negative.

```
if(MACD[stock]≥0){
signal=1
else
signal=0
}
```

3.2.5 Percentage Price Oscillator (PPO)

A sell signal is generated if the estimate of PPO is negative and vice versa.

```
if(PPO[stock]≥0){
signal=1
else
signal=0
}
```

3.2.6 Relative Strength Index (RSI)

RSI greater than 80 indicates an overbought level that suggests an offer signal and RSI less than 20 indicates an oversold level that suggests a bid signal. Heuristics are applied for the values of relative strength index in the range from 20 to 80 that if current period RSI is greater than past 10 period averages of RSI then a bid signal is given and vice versa.

```
if(RSI[stock]≥80){
//overbought
signal=0
}else if(RSI[stock]≤20){
//oversold
signal=1
}else if(RSI[stock]≥RSIMA[stock]){
//apply heuristics
signal=1
}else{
signal=0
}
```

3.2.7 Rate of Change (ROC)

A non-negative estimate of ROC is an indication to bid a stock and offer a stock otherwise.

```
if(ROC[stock]≥0){
signal=1
else
signal=0
}
```

3.2.8 Williams Percent R (WPR)

WPR oscillates on a negative scale that ranges from -100 to 0. Williams considered values less than -80 as oversold and more than -20 as overbought. Heuristics are applied to select an action for other values.

```
if(WPR[stock]≥-20){
//overbought
signal=0
else
signal=1
}

if(WPR[stock]≤-80 ){
//overslod
signal=1
else
signal=0
} else if (WPR[stock]≥WPRMA[stock]) {
//apply heuristics
signal=1
else
signal=0
}
```

All bid or offer signal bits generated using the illustrated technical indicators are combined together to create a vector of fixed length, equal to the number of indicators, representing the current market situation of stock determined using both short-term and long-term information that is passed as input to the stock trading agent which selects a best action to execute with the objective to earn optimum returns.

There are about hundreds of technical indicators but our current research scope is limited to nine indicators which are the most popular indicators used in wide for the analysis and also for the trading. Other reason for choosing these indicators as the candidate features for our trading model is because most of the technical indicators are derived using just the close price but our features are chosen in a manner that covers an entire technical information available that contains open price, close price, lowest price, highest price, and volume of shares traded.

3.3 Implementation

XCS is a rule based intelligent model that learns through experience using the reinforcement learning, genetic algorithms and heuristics. It maintains the knowledge of a problem domain in the form of a set of rules or classifiers called population set [P]. A rule set is generated at the beginning with random condition-action rules and default values are assigned to rule parameters such as prediction, prediction error, action set size, and fitness.

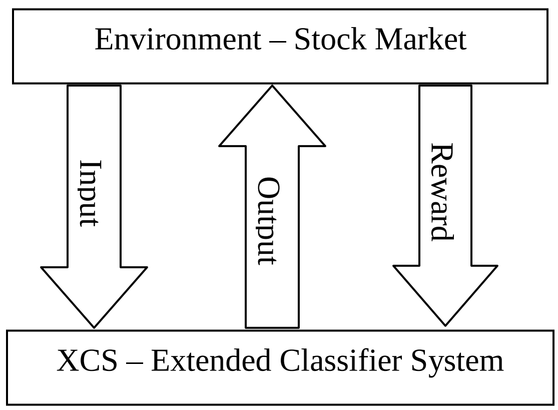

Figure 5 - Interaction of XCS with Environment

The above Figure 5 - Interaction of XCS with Environment illustrates an interaction of extended classifier system with the environment which is a stock market simulator in our case. The classifier system senses the current market situation in the form of a series of bits (0's and 1's) where each bit represents a proposed action or output or signal generated using one of the chosen indicators.

The nine technical indicators mentioned in first chapter are used to represent a major part of input to the XCS. The order of these indicators does not matter but all the bits should be laid in same order each time in order to learn and that order chosen for the representation of our input is SMA, EMA, MACD, PPO, CMF, ROC, CMI, RSI, and %R.

For example an input condition with the following bit string 101011010 indicates SMA, MACD, CMF, ROC, and RSI indicators suggests to bid on a share and other indicators such as EMA, PPO, CMI, and %R suggests to offer a share.

The input condition might be represented using continuous values from [0, 1] but the genetic algorithm is applicable for genomes represented as a string of bits. The learning of classifiers in a population set using continuous values would be much complex as well, so the research is just limited to non continuous values 0 and 1.

The input to XCS is also supplied with reward also known as payoff for all the possible actions along with the condition. In our model, we have chosen the actions for execution are either to bid or offer. The actions to execute are encoded as well using the bits 0 and 1 where 1 is used for the action to bid and 0 is used for the action to offer a share.

The payoff is a numerical value that represents a profit or loss incurred through the execution of a chosen action. For instance when a stock is bought for 10 bucks and sold for 12 bucks on immediate next period then the current reward upon offering a share would be 2 bucks. The positive values represent a profit and negative values represent a loss incurred.

E.g. The format of input to XCS, 101011010 $0 \rightarrow 2$ $1 \rightarrow -3$

The above instance states that market situation is represented using a bit vector 101011010 and upon executing an action to offer or sell a reward of 2 units is received otherwise a reward of -3 units is received from the environment using which the knowledge domain is modified. That is, if the agent proposes an action to bid then the model would receive a payoff of -3 units which is an indication of incorrect prediction and the agent learns using this so that it would predict a correct action from the next time in the similar market situation.

3.3.1 Random Generation of Individuals

The individuals or classifiers or rules of a population set are generated through condition-action pairs and other parameters of a classifier were assigned to default numerical values. The number of rules generated is a constant integer, 220, value known as population size.

The condition of a classifier is a string represented using the alphabets {0, 1, #} and the size of a condition is equal to the number of features used to represent a market situation (nine). The hash bit is used to indicate a do not care which means it does not matter whether it is either a bit 0 or a bit 1. The action of a classifier indicates the possible signals to be suggested for execution which are bid a stock encoded using a bit 0 and offer a stock encoded using a bit 1.

A likelihood of hashing $p_{\#}$ is taken for the inclusion of number of do not care bits in the condition, a higher value would generate a more general rule and vice versa. Other main parameters of a classifier are prediction, prediction error, accuracy, fitness, action set size.

Prediction indicates an expected reward when the classifier is selected. A small integral value, 5, is chosen as its default value.

Prediction error is a difference between the expected payoff and the actual payoff received from the environment. A small error, 0.00, is taken as its default value.

Fitness estimates the usefulness of a classifier. A small numeric value, 0.01, is selected as its default value to be considered during the random generation of individuals.

3.3.2 Formation of Match Set [M]

The input condition perceived from environment is matched with each classifier condition in the population set [P] and a set of rules formed is called as a match set [M].

In the formation of a match set a classifier bit 0 matches with 0 of input, 1 matches with 1 of input, and # matches with both 0 and 1 of an input.

A covering mechanism is applied when the input condition do not match with none of the classifiers condition or when the number of actions present in the match set is less than 2. A new rule is generated in the covering mechanism with a condition that matches to the input condition using 0s, 1s, #s and an action that does not exist in match set is selected for action. All other parameters of a classifier are assigned with the default values.

3.3.3 Total Predictions

The estimate $\sum_a pf/\sum_a f$ is computed on taking the ratio of the sum of the product of prediction and fitness values to the sum of fitness values for all classifiers that propose an action, a, in match set [M]. It is also known as the total or system prediction for the action, a, which is a measure of reward received for each action in the match set after its execution. Thus an action with the highest total prediction is chosen for execution according to an algorithmic description of Wilson's XCS (Martin & Stewart, 2000).

The values of total predictions for each action are stored in a linear array whose length is equal to the total number of possible actions for execution. The initial value is taken as zero for each new input and it is incremented after computing the total prediction for each action in match set.

3.3.4 Formation of Action Set [A]

The action set [A] is a subset of the match set that includes all classifiers from the match set that propose an action, a, which is suggested for execution. If an action, $a_{1,}$ has the highest total prediction then it is the proposed action according to XCS model and all the rules in a match set [M] that propose an action, $a_{1,}$ are taken to form a new classifier set known as action set [A].

Therefore a population set is the super set of both match set and action set, while match set is the super set of action set. [P] is superset of both [M] and [A] as [M] is derived from [P] and [A] is derived from [M]. Thus a relationship between the three classifier sets [P], [M], and [A] could be defines as follows $[A] \subseteq [M] \subseteq [P]$.

The rules or classifiers in action set are the active rules and hence are updated using the reward received for the proposed action from the environment or stock market simulator.

3.3.5 Updating Fitness

A classifier that belongs to a rule set, [P], is represented using a five tuple {condition, action, prediction, prediction error, and fitness}.

The following pseudo code derived from Wilson's XCS (Martin & Stewart, 2000) is used to update the fitness and other parameters of a classifier in action set.

//update prediction in [A]

prediction ← prediction + β × (actual payoff - prediction)

//update prediction error in [A]

prediction error ← prediction error + β × (|actual payoff - prediction| - prediction error)

The action set size of a rule in [A] is updated as well similar to prediction and prediction error.

//update fitness in [A]

accuracy sum←0

initialize accuracy vector κ for each classifier in [A]

if(prediction error<ϵ_0) κ ← 1

else κ ← α×(prediction error/ϵ_0)$^{-v}$

accuracy sum ← accuracy sum + κ

for each classifier in [A]

fitness ← fitness + $\beta \frac{\kappa}{\text{accuracy sum-fitness}}$

Where α, ϵ_0, and v are used to update the fitness and β is a learning rate used for updating the prediction, prediction error, action set size and fitness as well. The values for β, α, ϵ_0, and v are constant and the numerical values taken for these are β=0.2, α=0.1, $\epsilon_0 = 1$, and v=5 (Martin & Stewart, 2000).

3.3.6 Genetic Algorithm

In traditional learning classifier systems the GA is applied on the entire [P] but in the XCS algorithm the GA is applied just on the classifiers in action set, [A]. It is one of the reasons that XCS does perform better than other traditional LCSs apart from another reason that XCS updates fitness using accuracy of the reward predictions rather than just the reward.

GA is applied on action set after the execution of certain iterations or perceptions, θ_{GA}, of input from the environment. It is applied once after sensing θ_{GA} inputs from the environment. The θ_{GA} is constant and it is chosen as 70 after testing with the various values from 10 to 100. Therefore GA is applied after perceiving 70 inputs and the action set has atleast 2 classifiers to perform an evolution process using the genetic operators such as mutation, and one point crossover.

The likelihood of mutation for a condition string, 10%, is used to alter the number of bits in a classifier condition and it also alters the action bit with a 50% likelihood of mutation.

Unlike genetic mutation, the genetic crossover operation modifies just the classifier condition and keeps the action bit unaltered. The locus is chosen in random from one to nine to perform a single point crossover genetic operation on two parent individuals and the subparts are then interchanged to form two child individuals.

Before the application of GA, two parent classifiers are selected from the action set [A] and are cloned to perform genetic operations and evolve fittest classifiers. The roulette wheel selection approach is used to select the parent individuals that choose two rules with highest fitness. It then applies genetic operations, mutation and one point crossover, to evolve the fittest rules. The reproduced classifiers are then added to a population set [P] and then two rules with least fitness measure are deleted from the population set [P] using a roulette wheel deletion approach.

The other classifier parameters of child individuals such as prediction, prediction error, action set size, and fitness are the average estimate of parent individuals.

The framework of an extended classifier system implemented using the Wilsons XCS (Martin & Stewart, 2000) is represented as shown underneath in Figure 6 - Framework of XCS.

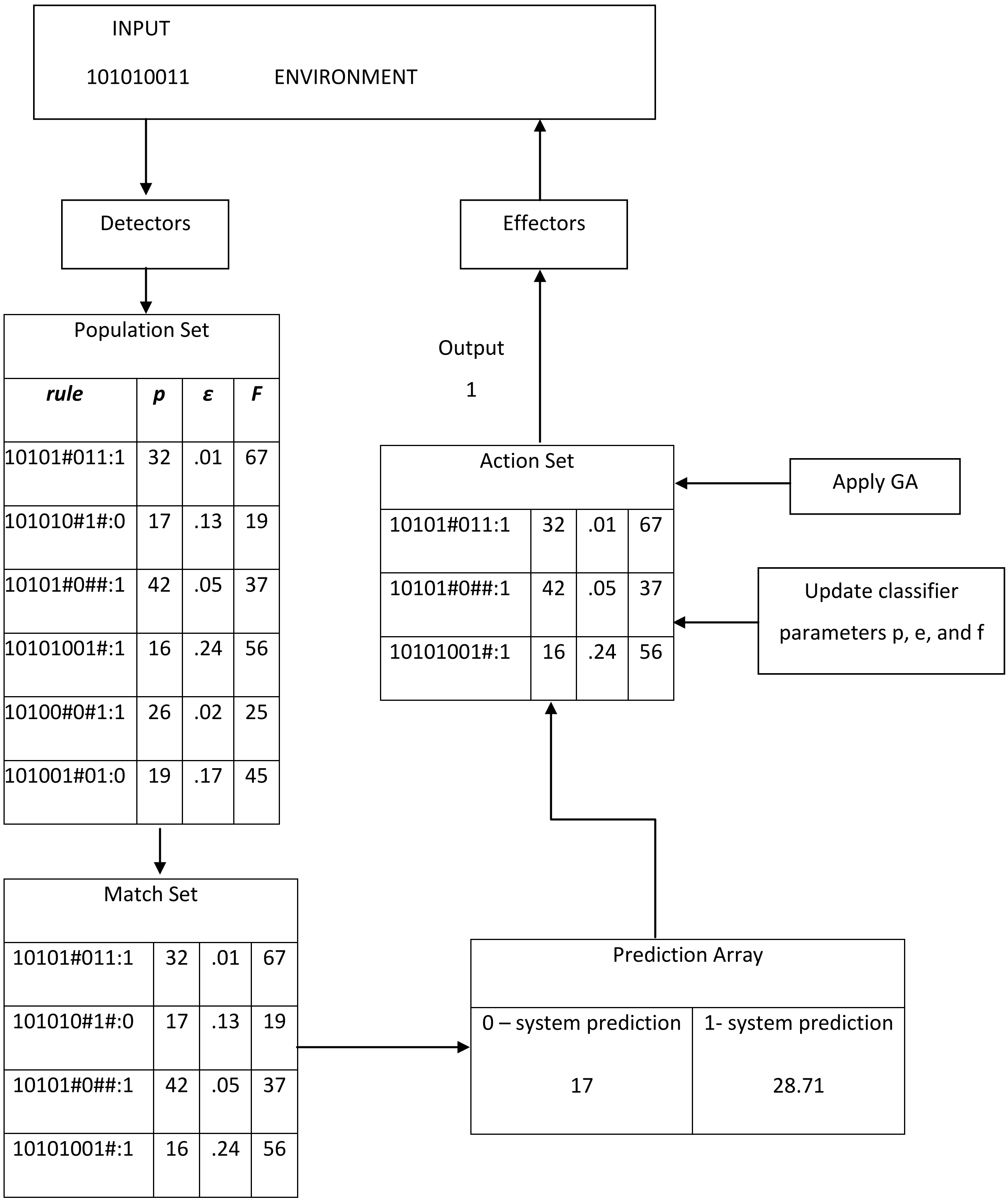

Figure 6 - Framework of XCS

Chapter 4 - Results and Evaluation Criteria

XCS agent generates a random set of classifiers known as population set [P] using the seed or psuedo random value. The [P] represents knowledge about a problem domain which has just the random individuals at the beginning but the rules keep on improving on perceiving more inputs from the environment. The seed value is used to genetare a different random individuals and test the performance of trading model. The results for realtime stocks have depicted that XCS model learns in a simialr manner even with different individuals at the beginning of a learning phase as there is not much difference between the final profits or returns.

The agent is tested using three real time stock data that consists of inputs simulating a stock market dataset of 10X12X4 = 480 weeks. XCS maintains a variable, bank account or total payoff or system payoff, to store the amount of profits received which is a summation of rewards earned from the environment.

XCS is tested for three different stocks that simulate an environment and the results are shown in a table and in a chart. The table on left hand side presents a pseudo random value, stock name and the corresponding wealth acquired after learning from the specific stock data. Table on the right hand side presents the intermediate results that indicate the amount of wealth acquired after scanning three hundred inputs from the environment.

Evaluation is done using the strength based mechanism where the trading model or agent performance is estimated using the sum of total rewards received so far and comparing it with the of total rewards received in the previous instances. From the results shown it is quite evident that XCS is learning from the stock market simulator in order to maximize the bank account that represents a total reward increasing along timeline which is an indication that the XCS agent is improving with experience.

Table 1 - INFY ADR

INFY ADR	
Seed	Total Returns
666	29978
777	31204
888	30560
444	29085

300	600	900	1200	1500	1800	2100	2400
1041	5039	9070	13299	17384	21598	25734	29978
1261	5397	9414	13704	18141	22442	26762	31204
1204	5428	9545	13710	17947	22096	26237	30560
1235	4941	8955	12916	16917	20743	24786	29085

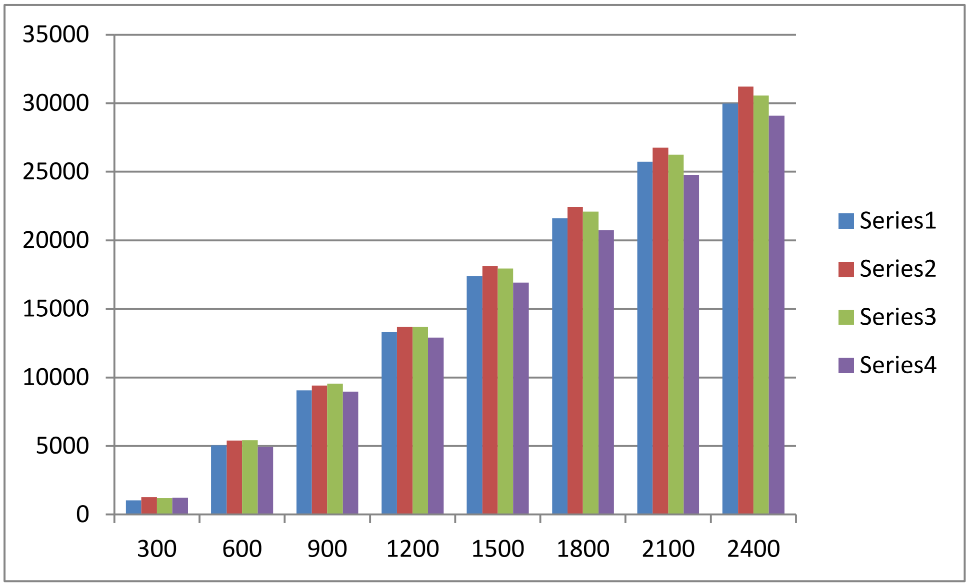

Figure 7 - INFY ADR

The left hand side of above Table 1 - INFY ADR presents the total optimum returns after the perception of all inputs from the environment while the right side depicts the optimum returns at regular time intervals, after every three hundred inputs.

The chart shown for INFOSYY ADR Figure 7 - INFY ADR is plotted with the timeline on horizontal axis and the profits on the vertical axis. It is actually the number of inputs that were plotted on the horizontal axis that even represents the timeline as the inputs ordered with the increasing time period.

The returns or profits are growing with experience which is directly evident from the results as the value of bank account is continuously rising with no fall at all. Also it is evident from the results that there is no quite difference in the total returns through generating different rules at the beginning of a learning phase.

Table 2 - CSC

CSC	
Seed	Total Returns
777	29564
600	29291
590	25802
500	26311

300	600	900	1200	1500	1800	2100	2400
1303	5076	9123	13053	17435	21450	25604	29564
1393	5371	9283	13285	17637	21538	25483	29291
1251	4668	8215	11590	15327	18728	22391	25802
1375	4930	8239	11752	15376	18926	22607	26336

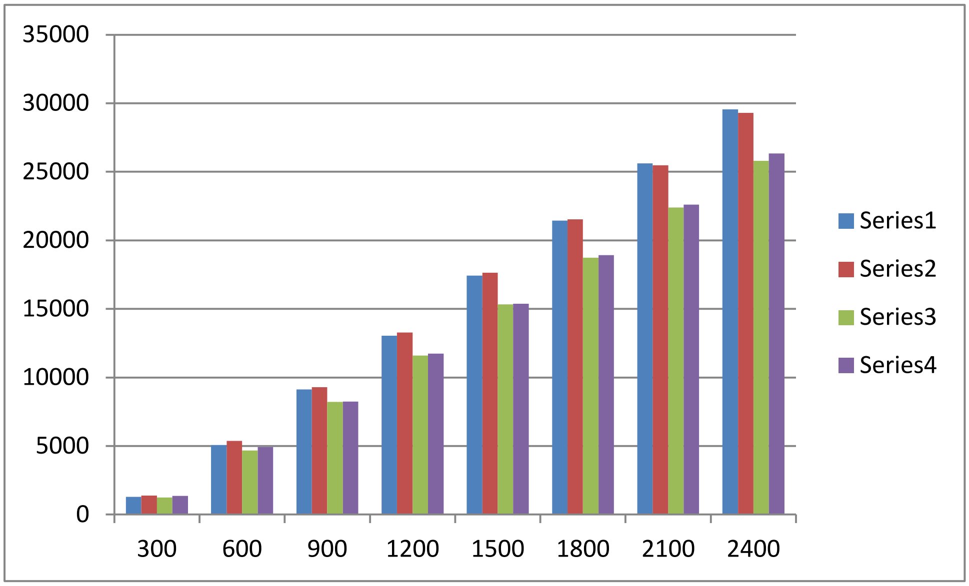

Figure 8 - CSC

The left hand side of above Table 2 - CSC presents the total optimum returns after the perception of all inputs from the environment while the right side depicts the optimum returns at regular time intervals, after every three hundred inputs.

The chart shown for stock Figure 8 - CSC is plotted with the timeline on horizontal axis and the profits on the vertical axis. It is actually the number of inputs that were plotted on the horizontal axis that even represents the timeline as the inputs ordered with the increasing time period. The returns or profits are growing with experience which is directly evident from the results as the value of bank account is continuously rising with no fall at all. Also it is evident from the results that there is no quite difference in the total returns through generating different rules at the beginning of a learning phase.

Table 3 - CTS

CTS	
Seed	Total Returns
500	45351
300	41973
200	43820
250	42407

300	600	900	1200	1500	1800	2100	2400
1920	7707	14020	20355	26495	32392	38821	45351
1671	7480	12994	18603	24150	30141	35853	41973
1777	7319	13515	19643	26421	32064	38081	43820
1825	7241	13653	19082	24927	30561	37160	42372

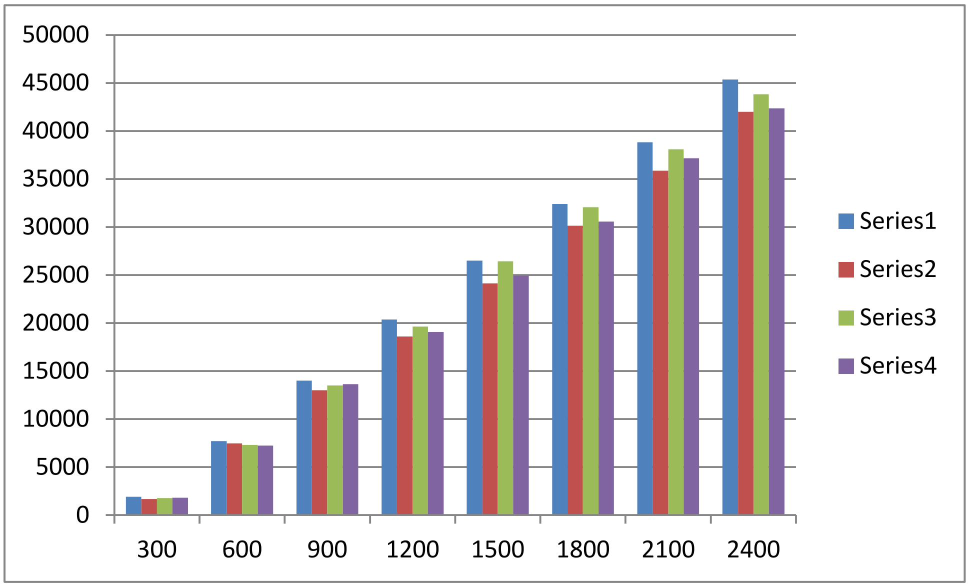

Figure 9 - CTS

The left hand side of above Table 3 - CTS presents the total optimum returns after the perception of all inputs from the environment while the right side depicts the optimum returns at regular time intervals, after every three hundred inputs.

The chart shown for stock Figure 9 - CTS is plotted with the timeline on horizontal axis and the profits on the vertical axis. It is actually the number of inputs that were plotted on the horizontal axis that even represents the timeline as the inputs ordered with the increasing time period.

The profits are growing with experience which is directly evident from the results as the value of returns is continuously rising with no fall at all. Also it is evident from the results that there is no quite difference in total returns through generating different rules at the beginning of a learning phase. Series 1, 2, 3, and 4 plotted in chart represents the profits at different pseudo random values as shown in the order of seeds in the left hand side figure.

Table 4 - Correlation

seed = 111, N=220			payoffs at regular intervals							
27 bits	total payoff		300	600	900	1200	1500	1800	2100	2400
CSC	43705		1873	8214	13818	19553	25084	31191	37571	43705
INFY	43029		2363	8377	14171	20066	26016	31741	35635	42029
CTS	43318		2024	8031	14050	20053	25649	34465	39352	46318

Table 5 – CSC

seed	population size	total returns	population size	total returns	population size	total returns	population size	total returns
111	150	42297	220	44161	250	44263	270	45454
222	150	38075	220	47415	250	41891	270	36112
333	150	40428	220	49918	250	49678	270	51420
444	150	41785	220	41408	250	43987	270	38807
		162585		182902		179819		171793

The results depicted in Table 4 - Correlation presents the profits earned for same stocks which is a bit higher than just taking an individual stock into consideration. The correlated stocks for CSC is a combination of three IT stocks CTS, SYNT, and CSC. In the same mannner for INFY ADR the correlated stocks chosen are CTS, SYNT, and INFY ADR. The correlated stocks considered for CTS stock are INFY, SYNT, and CTS as well because it has the perfect correlation with itself. Taking correlation of stock features performed better as the high correlated stocks behave in the same manner, that is if one rises then the other does with a little time gap and vice versa.

The results shown in Table 5 – CSC is the returns for CSC stock at different population sizes keeping the the pseudo random number as fixed. The results indicate that the agent is improving and providing optimum returns when the population size is fixed at 220. Thus, it is chosen as the default value for all previous experiments.

Chapter 5 - Conclusions and Future Work

The extended learning classifier system for stock trading is doing better as the total returns has never gone into the negative region which is an indication of incurring loss for the stock investor. The application of correlation of stocks as featutres in test bed has enhanced the performace of XCS trading agent resulted in a bit higher returns as compared to individual stock features. It is also an sign that there is still a scope of improvement in order to generate even higher returns to an investor. It is possible to improve the current trading model on considering the fundamental indicators, extending the length of an input condition and classifier condition.

The model could provide better results when other indicators such as Bollinger bands were also taken into consideration. Again it is not true that a higher number of indicators would give more returns. Therefore a precise selection of both the technical indicators and fundamental indicators would generate more profits for the investor.

The stock trading using XCS algorithm could generate even higher profits to the investor using an application of portfolio management on stocks. The XCS trading agent could be extended to achieve portfolio management where a trader invests in multiple stocks in order to minimize the loss in cases when the firm is bankrupted.

Our model is just suggesting the trader to either bid a stock or offer a stock but it is not informing the investor number of shares either to bid or offer. The extension of this would generate better returns upon maintaining the amount of wealth that a trader has acquired or lost after each auction on a stock. The action string could also be extended to have two bits so that an investor could be informed to hold a stock whenever there is no advantage in either bidding or offering a stock. It is required because sometimes there would be neither profit nor loss then it is not a good decision to trade in such situations as the investor would lose some amount to the stock broker in the form of a commission. It is unfeasible using the current action string because a single bit can support a maximum of two possible actions to execute.

References

Andrew Lo, W., & Mackinlay, A. C. (2002). *A Non-Random Walk Down Wall Street.* Princeton: Princeton University.

Bodie, Z., Kane, A., & Marcus, A. J. (2008). *Investements, Chapter 12, Market Efficiency.* McGraw-Hill Companies.

Booker, L. B., Goldberg, D. E., & Holland, J. H. (1989). *Classifier Systems and Genetic Algorithms.* Ann Arbor: The University of Michigan.

Burton, M. (1999). *A Random Walk Down Wall Street.* New York, NY, USA.

Button, R. S., & Barto, A. G. (1998). *Reinforcement Learning: An Introduction.* Cambridge: MIT Press.

Chauhan, A. (2008). *Automated Stock Trading and Portfolio Optimization Using XCS Trader and Technical Analysis.*

Chen, M. C., Lin, C. L., & Chen, A. P. (2007). *Constructing a dynamic stockportfolio decisionmaking assistance model: using the taiwan 50 Index constituentsas an example.*

Chen, P., & Chen, M. Y. (2005). *Integration extended classifier system and knowledge extraction model for financial investment prediction: An empirical study.* Taiwan: Institute of Information Management, National Chiao Tung University.

Choudhry, R., & Garg, K. (2008). *A Hybrid Machine Learning System for StockMarket Forecasting.*

Clarke, J., Jandik, T., & Mandelker, G. (1999). *The Efficient Markets Hypothesis.*

Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance* , 383-417.

Gershoff, M., & Schulenburg, S. (2007). *Collective behavior based hierarchical XCS." Proceedings of the 2007 GECCO conference companion on Genetic and evolutionary computation.*

Hellstrom, T., & Holmstrom, K. (1998). *Predicting the Stock Market.*

Holmes, J. H., Lanzi, P. L., Stolzmann, W., & Wilson, S. W. (2002). Learning classifier systems: New models, successful applications. *Information Processing Letters* , 23–30.

Kalyvas, E. (2001). *Using Neural Networks and Genetic Algorithms to Predict Stock Market Returns.*

Kendall, M. (1953). The Analysis of Economic Time series Part I: Prices. *Journal of the Royal Statistical Society* .

Malkiel, B. G. (2003). The Efficient Market Hypothesis and Its Critics. *Journal of Economic Perspectives* , 59-82.

Mandelbrot, B. (1966). forecasts of future prices, unbiased markets and martingale models. *Journal of Business* , 242-255.

Martin, B. V., & Stewart, W. W. (2000). An Algorithmic Description of XCS.

S, S., & P, R. (2001). Explorations in LCS Models of Stock Trading. *Advances in Learning Classifier Systems* , 151-180.

Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review* , 41-49.

Shah, V. H. (2007). *Machine Learning Techniques for Stock Prediction.*

Sigaud, O., & Wilson, S. W. (2007). *Learning Classifier Systems: A Survey*.

Smith, S. F. (1980). *A Learning System Based on Genetic Algorithms.* Pittsburg: Department of Computer Science, University of Pittsburg.

stockcharts.com. (2011). Retrieved September 2011, from http://stockcharts.com/

Stone, C., & Bull, L. (2004). *Foreign Exchange Trading using a Learning Classifier System.* Bristol United Kingdom: niversity of the West of England Bristol.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning.* Cambridge: MIT Press.

Tsaih, R. H. (1998). Forecasting S&P 500 stock index. *Decision Support Systems* , 161–174.