I don’t care whose fault it is!

Or

An introduction to the

Short-Term Forecasting Theory,

implementing fuzzy-logic and neural networks.

Written By

Jordan Bernstein, 2006

I don’t care whose fault it is!
Or an introduction to the
Short-Term Forecasting Theory,
implementing fuzzy-logic and neural networks.

by Jordan Bernstein
at the University of Waterloo,
Independent Studies Program

Printed in Canada

Thank you,

Thank you to each of my thesis advisors, Ken Vetzal and Alonso Perez, who always tolerated my numerous setbacks, and who each acted as my ever-faithful driving instructors on this big-rig of a paper. Thank you for not letting me crash!

To my undergraduate advisor Anne Dagg, and everyone in the Independent Studies department, thank you for your never-ending enthusiasm and support. You have no idea how much you've helped!

To my friends, and my family, thank you for pushing! I often pushed back, but you could always push harder. Through hard times and late nights, your support kept me going. I could have never done this without you.

Jess my love, thank you for teaching me that it's better to touch the moon than to always be reaching for the stars!

To Zaidy.

I did it!

Table of Contents

i. Executive Summary

In contradiction with much conventional economic theory, this thesis argues that successful short-term forecasting is both possible and practicable. Beginning with the assumption, and widely-held belief, that there are patterns to be discovered in the stock market, the thesis develops the Short-Term Forecasting Theory (STFT) to demonstrate how useful and accurate short-term forecasts might be achieved. In short, this thesis posits that, if short-term financial forecasting of an equity can be broken down to a mechanical procedure, the problem of short-term forecasting is reduced to the question of finding the proper tools for this procedure. This thesis presents two computing methods – fuzzy logic and neural networks – that, when combined, could serve as an appropriate tool for implementation.

ii. Preface

"I don't care whose fault it is!"

Most of us can recall, at least one (and often many) authoritative figures in our childhood using that statement after some sort of skirmish with our peers. No matter how hard we tried, how much we cried, or how loud we yelled, it just didn't matter whose fault it was.

"Outrageous!" I must have thought to myself at the tender age of five, when first confronted with this seemingly impassable statement of fact. "How is it that it doesn't matter 'whose fault it is'? A civilized society is based on fault! Fault is responsibility, and from responsibility grows progress and reward. Newton tells us, that for every action there is an equal and opposite reaction. Without fault, there is no reaction, and with that, the wheels of production cease to move. If we don't have fault in our society, how do we know who to punish and who to reward for telling on them?"

Right or wrong, they may have had something. In terms of forecasting, why do we care who started it or whose fault it is? Does it really matter *why* the price of oil has risen 25% over the past few hours? No, not if we know that it will directly result in the price of plastics to rise 50% in the next four hours! This is the basic premise of this paper. A premise that, while appearing juvenile in its simplicity, actually proves to be quite complicated and perhaps flies in the face of years of economic theory.

The progression of this paper has been a long and tortuous one before it became the Jewel of the Nile you see before you.

It started out as a simple premise that if Fuzzy-logic is good for pattern recognition, and Neural Networks are good at learning things, it should be a simple matter of bringing the two together in order to forecast the stock markets of today. This proved to be more difficult than anticipated!

I then turned to the idea of explaining why fuzzy logic and neural networks should be used for financial forecasting. While I maintain that the combined use of fuzzy-logic and neural networks is a viable technique for short-term financial forecasting, I soon realized that it wasn't the methodology with which most, if not all, scholars were concerned, but rather that short-term forecasting may not even be possible.

Only then did I realize that the purpose of this paper was to explain the painfully simple premise that it really doesn't matter *whose fault it is*!

*****

1. Introduction

Long has it been since someone first tried to see into the future, usually by praying to a divine entity or perhaps through the use of a crystal ball, hoping to benefit in some way or another. Today, however, our crystal balls are in the form of computers that are capable of performing millions of small tasks in virtually no time, and the benefit is money. This thesis does not claim to have developed a new type of crystal ball, nor does it claim to be able to make anyone rich by reading it. What it does claim is a theory that demonstrates that short-term forecasts are likely to be successful – a theory that, if correct, contradicts many accepted economic theories. Along with this theory are presented two computing methods that, when combined, could serve as an ideal tool for implementation.

This thesis suggests that short-term financial forecasting of a selected equity is both possible and practical and attempts to explain why by introducing the Short-Term Forecasting Theory (STFT). The STFT is a theory that, while not exclusive to *financial forecasting*, arguably lends itself well to that task. The term *financial forecasting* has been defined by the Accounting Standard Executive Committee of the American Institute of Certified Public Accountants (AICPA) as

> …an estimate of the most probable financial position, results of operations, and changes in financial position for one or more future periods…based on management's judgment of the most likely set of conditions and most likely course of action.
> (Giroux and Rose 1981:1)

For the purpose of this paper, the definition has been expanded to: *Any speculative forward-looking process regarding the price of a single publicly traded equity.*

While it is simply impossible to visit the future and learn of things to come, it is entirely possible to predict the future course of many occurrences based on their prior performances. It would appear that this would hold true for future stock prices, at least in the very short-term (Colby, 2003:iii). The term *short-term* has many different meaning to different people. To a long-term investor who might hold a stock for 10 years, a one-year investment would be considered short-term. However, in this thesis short-term is considered to be from one minute up to a few hours. Another term used in this thesis is *immediate-term,* and this should be taken to mean a time horizon that is less than one minute. This paper primarily deals with the functions of publicly held, individual corporate stock and the stock markets in which they trade. In short, the premise of this thesis is that, if short-term financial forecasting of an equity can be broken down to a mechanical procedure (Section 3), the problem of short-term forecasting is simplified into finding the proper tools for the procedure. The appropriate tools are argued in this thesis to be fuzzy-logic (Section 4.1) and neural networks (Section 4.2)

Section 2 gives a brief overview of the history and common theories concerning financial forecasting. Section 3 explains why it should be possible to forecast short-term financial data based on the assumption, and widely-held belief, that there are patterns to be discovered in the stock market. In Section 4, the tools for forecasting are presented: fuzzy-logic and neural networks, in Section 4.1 and Section 4.2 respectively. Section 4.3 explains how combining the different technologies found in Sections 3, 4.1, and 4.2 can be used to formulate a practical forecasting mechanism. Section 5 summarizes the different theories and tools presented in this thesis.

*****

2. History of Financial Forecasting

Numerous methods of financial forecasting exist today, each with their own benefits and drawbacks. Four popular methods are described in the following subsections. Unfortunately, none of these more popular forecasting techniques is fully suitable for short and immediate-term forecasting of ever-evolving trading markets. Traditionally, there are three classifications of forecasting techniques. There are *qualitative* methods, *time series analysis* methods, and *causal* methods. Qualitative Methods rely upon the judgment and experience of bank management and other trained observers. Time Series Analysis relies entirely on patterns in historical data. Causal Methods express significant relationships mathematically. Because of its ability to equate key variables, Chambers, Mullick, and Smith consider causal forecasting models the most sophisticated of all forecasting tools (Giroux, *et al* 1981:44).

Each of the three categories encompasses several different techniques. For example, both the popular Delphi[1] and Market Research[2] methods fall under the Qualitative Methods. Moving Average[3], Exponential Smoothing[4], and the sophisticated Box-Jenkins[5] methods all belong to the time series analysis

[1] An iterative process utilizing a panel of experts to achieve a consensus. Typically used for long-range planning (Giroux, *et al* 1981:44).
2 Systematic techniques for testing current market situations. Typical applications include long-range planning and new product sales (Giroux, *et al* 1981:44).
3 Weighted average of a number of consecutive points in time (Giroux, *et al* 1981:44).
4 Similar to moving averages, but more recent data points are given greater weight (Giroux, *et al* 1981:44).
5 Time series data is fitted with a mathematical model (which must be identified) that limits forecast error to the greatest extent possible (O’Donovan, 1983:15).

approach. Causal Methods include both Regression Analysis[6] and Econometric Models[7]. These are to name just a few of the more popular methods in use today in the banking industry.

*****

2.1 The Delphi Method

According to Greek mythology, those interested in learning what the future held in store for them would seek the counsel of the Delphic oracle. The Rand Corporation developed a forecasting system suitably dubbed the Delphi method, which "systematically enlists the opinions of several experts, producing a single decision or consensus" (Greenberg, Baron, Sales, and Owen 2000:328).

The implementation of the Delphi method is relatively systematic, in that there is a very simple, well-defined, method of application. A question is put forth to a panel of experts, often by mail or internal memo, thereby eliminating the need for face-to-face meetings. Each expert individually responds, in writing, to the issue at hand. This information is then passed on to a project leader who compiles and reviews the responses. The project leader then redistributes all responses to each of the experts in the panel. The experts then, again in writing, comment and respond to arguments made by other panel experts. The project leader, once again, compiles and reviews the responses, this time looking for a consensus among the experts. If a consensus is reached, the conclusion can be implemented; however, if no consensus exists the process of redistributing all the responses to each of the experts, in order that they may comment, is repeated (Greenberg, *et al* 2000:328-329). Figure 2-1-1 represents the how the Delphi Method is applied.

6 Relates independent variable(s) to a dependent variable to be forecast by estimating an equation using least squares techniques (Giroux, *et al* 1981:52-53).

7 A system of interdependent regression equations describing one or more economic sectors (O'Donovan, 1983:5).

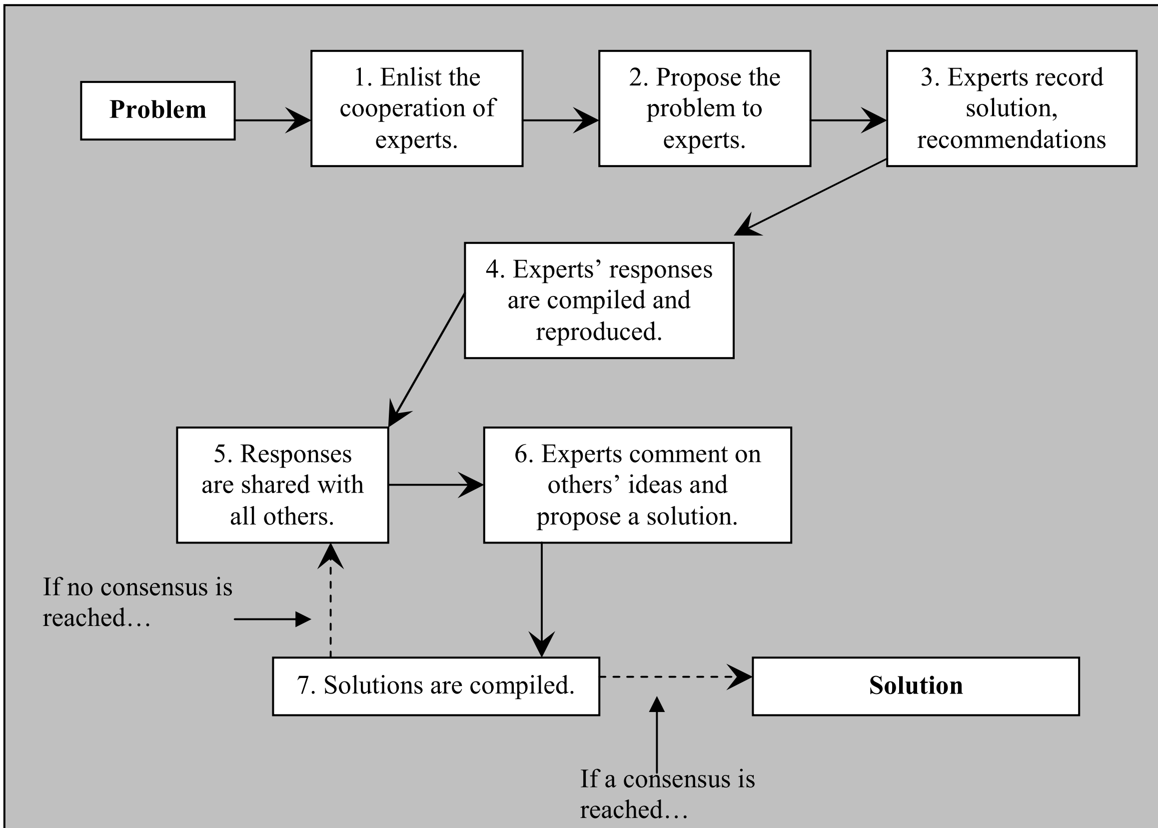

Figure 2-1-1 The Delphi Method–adapted from Greenberg *et al*, 2000:328

One of the advantages to this method is its ability to assemble a large wealth of knowledge and arrive at a single conclusion or range of possibilities. It also allows for the participation of remote parties. The absence of face-to-face meetings allows group members to participate from any location. It also tends to eliminate or reduce the *Groupthink*[8] phenomenon.

The Delphi technique is not, however, without its share of drawbacks. The most prominent of these is its tendency to be very time-consuming:

> "Sending out letters, waiting for everyone to respond, transcribing and disseminating the responses, and repeating the process until a

[8] "…term to refer to the tendency of some groups who work together over a period of time to produce poorly reasoned decisions. When social pressures and conflict-avoidance overtake the desire to rigorously question alternatives, contradicting opinions are suppressed in favor of concurrence seeking. Eventually, the growing cohesiveness within the group leads to a shared illusion of invulnerability." (Bazerman 1998:150)

> consensus is reached can take quite a long time. Experts have estimated that the minimum time required to use the Delphi technique would be more than 44 days…Obviously, the Delphi approach would not be appropriate for making decisions in crisis situations or whenever else time is of the essence." (Greenberg *et al* 2000:330)

Although the Delphi technique is often cited and used in financial forecasting, it is impractical for use as a technique for short-term forecasting due to the length of time necessary for proper execution.

*****

2.2 The Box-Jenkins Method

The Box-Jenkins method is considered by some to be among the most sophisticated and accurate time-series techniques. Unfortunately, with its sophistication come a great number of complexities and consequently increased costs (Giroux, *et al* 1981:59). The Box-Jenkins method involves comparing historic data with known forecasting models and determining which class of model best fits the data the user is trying to forecast. The model types used in the Box-Jenkins method are known as Autoregressive Moving Average, or ARMA models. The determination of which class of model to use is by nature complex and takes a great deal of skill and experience. Even the most skilled forecasters may arrive at different models given the same data set. The process of determining which model is to be used is in no way automatic and therefore the forecast is subject to the bias of the forecaster (O'Donovan 1983:11).

The Box-Jenkins approach may be divided into three fundamental stages, as outlined below (see Figure 2-2-1):

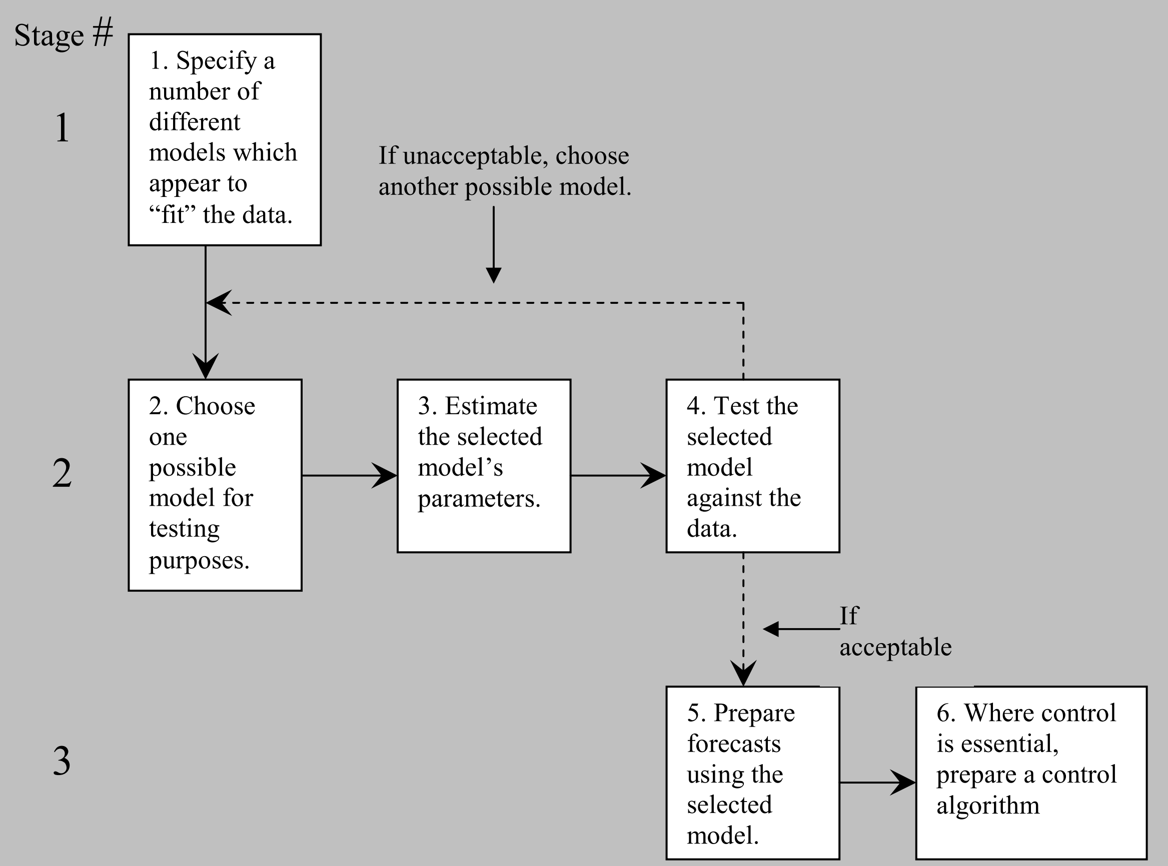

Figure 2-2-1 The Box-Jenkins Method

1. Selecting a suitable class of models for fitting to the observed time series[9].
2. Fitting an appropriate model from this class to the observed time series.
3. Using the fitted model to make forecasts of future values of the time series.

(O'Donovan 1983:77)

This method, as opposed to the Delphi method, is suitable for short-term forecasting. Once a proper model is selected and implemented, as long as there is no need for any model modification, forecasts can be made for any time horizon. Long-term financial forecasting is, however, not appropriate with the

[9]"A *time series* is a time-ordered sequence of observations of a variable. A *time-ordered sequence* is said to be one in which the observations are arranged in the order in which they were observed." (O'Donovan 1983:15)

Box-Jenkins method as markets do change and models therefore need modification. The Box-Jenkins method relies on the model being accurate and current. Any class of model, or individual model, cannot compensate for major market movement due to unpredictable circumstances, such as corporate mergers and changes in government policy. In order to compensate for some of its drawbacks the Box-Jenkins method is often used in conjunction with econometrics.

*****

2.3 Econometrics

Econometrics is the practical use of statistics to test economic hypotheses. Alternately put, "econometrics is the branch of economics dealing with the quantitative analysis of economic behavior" (Kelejian and Oates 1989:1-2). The use of statistics comes into play when there is a need to test a single variable with all other variables left unchanged. This can be done artificially through the use of statistics. Statistics allow us to estimate effects on a single variable by compensating for the movements of other variables. Two central issues of econometrics are the development of effective techniques to estimate the relationships of economic variables (Kelejian *et al* 1989:25), and the advantages of its ability to analyze complex interdependencies between economic variables (Giroux, *et al* 1981:57). A combination of econometrics and the Box-Jenkins method allows for a more accurate model that can better account for market fluctuations, and thus address one of the weaknesses of the Box-Jenkins method. It should be noted that "econometric models are highly useful for understanding the economic system and for the analysis of alternative management decisions and economic scenarios," (Giroux, *et al* 1981:58) and hence partially compensate for the weaknesses of the Box-Jenkins method employed alone. However, the use of econometrics is generally suited to long-term general market directions, rather than individual stock prices, and so, even in conjunction with the Box-

Jenkins method, does not lend itself well to short and immediate-term forecasting.

*****

2.4 Technical Analysis

Some would argue that technical analysis is as much an art as it is a science, and does not clearly fall into any one of the three categories listed at the start of this section – qualitative, time series, and causal. In brief, technical analysis is the systematic study of stock charts, most simply in the form of price vs. time, for known *market indicators* that suggest a potential trend or change in trend, one stock at a time. These market indicators, appearing as special shapes on price charts, indicate future prices based solely on historic similarities. Technical market indicators have been used since their creation in the mid 17^{th} century by Munehisa Homma, a wealthy Japanese rice trader, to profit from changes in the price of stocks or commodities. Technical analysis works by studying current price charts and searching for patterns that have historically preceded a known change, thereby signaling a potentially similar change. When a pattern is found, the current data is extrapolated to resemble the historic data, and a decision to buy, sell, or hold the particular stock in question is made based on this comparison. The most common, and simplest to understand, of the technical analysis charting techniques is the *candlestick* chart. A candlestick chart has a marker, known as a candlestick, every designated time period – one every month, week, day, hour, minute, second, or any other period that is deemed appropriate. Each candlestick graphically contains five pieces of information for its period: high price, low price, opening price, closing price, and overall gain or loss. When investors compare two or more sequential candlesticks, they look for known market indicators for clues as to when to buy, sell, or hold, that particular stock. Many other charting techniques exist under the technical analysis umbrella, however they each involve comparing a piece or group of data to historical data. Often these data are calculated though various types of averages or weighted averages, but always compare current data to historical data, in hopes of finding a pattern.

This method of investing is very different to the more traditional *fundamental* form of investing, where investors are primarily concerned with the financial viability of the company issuing the stock in question, and its relative performance to other companies with similar, competing, or related offerings. Technical analysts subscribe to the theory that markets are inefficient[10] and that markets move in trends, regardless of the causation of that trend (Colby 2002:5). While fundamental investors implicitly question market efficiency, as current prices would already reflect all available information, for the most part technical analysts do not concern themselves, as do fundamental investors, with any traditional profitability indicators.

The existence of the field of technical analysis greatly strengthens the argument that there are patterns in the financial marketplace in the short term and that if one could identify those patterns, financial forecasting becomes a technical issue, not a theoretical one.

[10] An *inefficient* market, as opposed to an *efficient* market (see 'Efficient Market Hypothesis' in Section 3.1), is not an expression describing how well a market is performing, but rather, is a theory suggesting that a market can display changes in equity prices that can be attributed to, not only new information, but also anything at all. An inefficient market allows for the possibility of changes to equity prices based solely on influences such as fear, greed and personal disposition. Inefficient investors lead to an inefficient market (Kingdon, 1997:5)

3. The Short-Term Forecasting Theory

Theories regarding short-term stock investments are not a new idea. The entire field of technical analysis, as discussed in Section 2.4, is dedicated to short-term stock trading for profit. Each method has its supporters and its opponents, each citing various pros and cons, as discussed in Section 3.1 and throughout this work. This thesis proposes the Short-Term Forecasting Theory (STFT) as a method for understanding the short-term predictability of relationships of seemingly unrelated variables. Many of the issues encountered with other similar theories are dealt with in the following pages, in an attempt to develop an accepted theory of short-term forecasting.

The proposed definition of the STFT is as follows:

> It is theoretically possible for **immediate** and **short-term** predictions to be made using historical time-series data for the reason that if a pattern of events exists between multiple specimens, it is unnecessary to know the initial cause of the pattern in order to exploit it, using pattern recognition systems. It is possible to make immediate and short-term predictions based on historical data, as pattern recognition is based on the *results* of the pattern instigator, not the instigator itself. *Therefore, the instigator need not be known; only the initial results.*

The subsequent sub-sections will answer the following five questions:

1. What is the problem with traditional forecasting?
2. What makes the STFT possible?
3. How does the STFT overcome the problems of traditional forecasting?
4. What are some of the limitations of the STFT?
5. What are the requirements of the STFT?

*****

3.1 What is the problem with traditional forecasting?

It is impossible to reliably predict all the future actions of any company, or of anything for that matter. Traditionally, stock forecasting relies on extensive market research and much educated guessing as to what a company and the general market will do in the future. The Delphi method, as discussed earlier, is often an element of this type of forecasting. The complex data produced from this type of research is of such a nature that it is impossible to break it down, categorically, to finite numbers and formulae. The mechanization of such qualitative methods seems, at this point in time, impossible. One might conclude, then, that any attempts to mechanize stock forecasting would be futile.

Moreover, conventional economic theory teaches that such efforts cannot yield valuable insights. The Efficient Market Hypothesis (EMH) implies that the prices of a stock will not change until new information, real or perceived, is available. When new information is made available, it is disseminated at such a quick rate as to render it impossible for an individual to have a meaningful advantage over anyone else (Kingdon 1997:5). John Maynard Keynes, the father of modern economics, noted that no one knows for certain what will influence the profitability and success of a particular company, and therefore, unless accurate predictions can be made about such things as every-day business activity, war, environment, and depression, it would seem impossible to consistently and accurately predict stock prices (Malkiel 1999:31).

Yet, the STFT claims that it is possible to make accurate short-term forecasts. The key to this advance is the theory’s ability to not only limit the data needed to make a successful forecast, but also its ability to easily incorporate other highly relevant data.

The advantage of such a forecasting system is obvious, as forecasting, whether short-term or long-term, must be performed predominantly by well-trained experts. Experts are the only users able to take advantage of many advanced forecasting techniques, due to their complexity. Each technique is unique and requires new training, understanding, and experience to be able to extract meaningful data. This need for training can be extremely time-consuming and costly. Not only does it cost to train an expert, but also, there is a cost associated with removing an expert from a productive position and rendering him/her useless, in a production sense, for the time needed for training. This weakness applies not only to those experts absent due to training. It also applies to those absent due to illness, vacation, turnover, maternity, weekends, and evenings. In fact, automation could greatly increase the potential for profitability on this basis alone.

All common forecasting techniques require very regular supervision, even after implementation. This supervision, in turn, adds to the overall cost of forecasting. An automated system would allow for the removal of the experts needed for implementation and regular supervision. The mechanization of the STFT would allow for such an automated system.

Cost is also an issue when considering implementing a computer aided forecasting system. Many systems require very high-end hardware to function properly, and all systems require costly, specially developed software. Maintenance and operation of these systems may also be costly and may often undermine the profitability of the system itself. In addition, there are currently no software products available that allow for complete, or near-complete, reliable automation, and while it might be argued that complete reliable automation is impossible, any amount of reliable automation would lead to a greater potential for profitability.

Many of the financial forecasting techniques used today are financial-specific, and cannot be transferred to other applications. This appears to be especially true of the short and immediate-term forecasting techniques. This lack of transference indicates a flawed or highly specialized theory. All generally accepted theories in physics, mathematics, and science must always hold true for any situation, or at least be adaptable within a certain realm of experimentation – social-science and behavioral theories must hold true, for at least, similar social circumstances. For example, Einstein's theory of relativity must hold true for any situation at any point in time and space, not just under controlled circumstances. An example:

$$\left.\begin{array}{r} 2+2=4 \\ \text{and} \\ 2\times 2=4 \end{array}\right\} \quad \text{One might then conclude that } x+x=x\times x \text{ or } 2x=x^2$$

In the situation where $x=2,$ the above equation works. However, the equation is clearly incorrect in general. This simplified example serves to show how a formula that appears to work and be correct in one context, may not necessarily be universally so. Generalized social theories may not necessarily be applicable within every culture and time period, however many similarities exist throughout the world, and within each of the specialized areas concerned, patterns may exist, and could therefore be recognized and analyzed. The STFT simply claims that if there are patterns to be found, it is possible to exploit them; it, in no way, says that all patterns are applicable universally.

The subject of financial forecasting is both one of pure-science and social-science. The addition of the social aspect is an immeasurably more complex one that, generally, lends itself to descriptive and predictive statistical theories, not purely mathematical ones. However, using statistic models it can be seen that there are, indeed, patterns to human behaviour, and not simply chaotic data. These are the patterns that the STFT exploits.

The complexity of all examined short-term forecasting techniques is very great. This, in turn, also limits the possibility of transference. Due to their highly complex nature, transference is limited because of the need to adapt and re-work the highly complicated technique to the new situation.

Of the commonly used short-term time-series forecasting techniques, all use a static model database. A static model database, or static knowledge-base, may be of relatively little use when concerned with stock markets. The collective and formal forums used to trade and describe the interrelated prices of individual stocks, known as stock markets, seem to be constantly changing over the long-term. Just as people's taste for fashion changes from one season to the next, so do investor's taste for their preferred area and method of investing. While commodities may be in vogue one year, over time, they may easily be replaced by newer, more exciting, technology sectors; risky gold ventures one year, stable "Blue-Chip" stocks the next. A static knowledge-base will not perform well when the model in use is no longer representative of the market. For this reason, a dynamic model or knowledge-base would clearly be superior to the equivalent static version (Kingdon, 1997:3). Using a static knowledge-base should result in a decreasing Rate Of Return (ROR) over time. As the market changes, the models used might not be able maintain an accurate forecast and will, consequently, decrease in performance. Depending on the rate of change in the market, the degradation of the static model will vary in acceleration. This degradation would cause an unpredictable ROR, and result in inconsistent forecasting ability. The consequence of an unpredictable ROR limits the practicality of static knowledge-based forecasting.

In financial forecasting, there is a general trend towards medium and long-term forecasting. Medium and long-term forecasting is not, and presumably cannot be, computerized or automated. Forecasters claim that a computerized long-term forecasting system is far too complicated to be a possibility. Certain aspects of long and medium-term forecasting may be computerized, but mainly rely on experience, educated guessing, and intuition. Neither guessing nor intuition may be programmed into a computer. They are non-logical operations, not compatible with a computer's logic processing. Long and medium-term financial forecasters maintain that their job cannot be performed by computers and, therefore, must be done by specially trained forecasters, such as themselves. Training is available, and is highly regarded, for long-term forecasting.

Another reason for the apparent lack of literature regarding short-term financial forecasting is that conventional economic views maintain that it is not possible to make accurate short-term financial predictions. The Efficient Markets Hypothesis, as mentioned earlier in this section, disavows the possibility of consistent, non-random, accurate predictions concerning the price of a stock. The EMH claims that unless privileged information is available, it is impossible to profit from a stock forecast because by the time one forecaster has all the necessary information to make a stock prediction, every other forecaster has the same information and has come up with the same forecast. Due to the speed with which the markets operate, it would be impossible to make a profit due to swift market transactions. An abundance of work has concerned itself with the implications of the EMH being true for all situations. The view of the traditional economist is that any gains, which might appear from any short-term financial forecasting technique, are the results of market randomness working in the favour of the forecaster, rather than successful market prediction.

Another argument against automated financial forecasting, especially that of a short-term nature, presented by those who follow the traditional economic view, is that financial markets are highly complex -- so much so, that it would be absolutely impossible to account for all variables in an economic system. Any attempts to produce an automated forecasting system will never be able to incorporate all possible variables and must, in the view of the traditional economist, inevitably fail.

The lack of academic interest in short-term financial forecasting, presumably due to its inherent risk and theories such as the EMH, can be seen by the deficiency of literature available, compared to that available for long-term financial forecasting. There is interest in short-term financial forecasting in the banking industry, but unfortunately for academia, most research into the subject is held in strict secrecy in the hopes that the exclusive use of a technology will result in a greater profit. This is because the greater the number of people creating the same forecast, the greater the pressure placed on the market to adjust, rendering the forecast progressively ineffective. For a more in-depth discussion of this issue, please see Section 3.4.

*****

3.2 What makes the STFT possible?

It must first be pointed out that successful forecasting, of any kind, usually predicts the actions of large groups, not individuals. Technical analysis is often criticized due to its inability to incorporate data external to the particular stock being studied, whereas the STFT incorporates data from *all* stocks in order to make a forecast. The STFT is a valid theory of forecasting because of its ability to predict the actions of large groups. With this in mind, it is appropriate to refer to a discipline where the actions of large groups, such as investment markets, are studied. A discipline which conforms to these criteria is commonly referred to as *crowd sociology*. Crowd sociology attempts to define similar actions in similar groups. By discovering similar actions within similar groups, we may attempt to predict the actions of these groups in similar situations, such as investors confronted with various market fluctuations. These large groups, often called crowds, share several common features.

> Members are uncertain about how to behave and what to expect. A feeling of urgency grips them, a feeling that something must be done *now*. The interaction of the emergent crowd produces a common mood or imagery and a conception of some kind of

> appropriate action. There is pressure on individuals to conform to this developing image or mood. The crowd is permissive; persons feel free to express attitudes that are ordinarily repressed, and express them more emotionally. They may also evolve new ways of acting and new attitudes. (Biesanz and Biesanz 1969:447)

The preceding analysis can be broken down into several statements relevant to the study of financial markets. Members of an investing market feel “uncertain about how to behave and what to expect,” and therefore cannot be expected to act in a rational manner in the short term. This corresponds to the notion that “all price movements have one thing in common: they are a reflection of the trend in the hopes, fears, knowledge, optimism, and greed of the investing public” (Pring 1991:5). These are not rational emotional foundations upon which to develop an accurate forecast, and yet, these emotions direct the price of a stock, regardless of rationality. “In [investors’] frenzy for money, market participants throw over firm foundations of value for the dubious but thrilling assumption that they too can make a killing by building castles in the air,” (Malkiel 1999:35). There is also a tendency to act immediately, as “something must be done now.” This does not allow for full analytical judgment and tends to produce more of a ‘knee-jerk reaction’ type of market execution.

The tendency of such non-rational crowd behaviour within financial markets perpetuates the investors’ predictions, valid or invalid, increasing the likelihood of a meaningful forecast. Because of the self-assimilating nature of a large percentage of investors, the forecast become self-fulfilling. Pressure on individuals within a large group further enforces a common mentality, perpetuating the disposition of the forecasters. When investors believe the price of a stock should go up, they are willing to pay a premium for the ability to buy while the stock is still, what they believe to be, reasonably priced. As more and more investors buy with the same mentality, the price of the stock they predicted would rise, does indeed rise, but not because the intrinsic value of the stock has risen, but rather, because the perceived value has risen and their earlier predictions are now confirmed.

Individuals of a group are also more likely to be emotionally persuaded, and will tend to base their actions on emotions such as fear or greed. The changing attitudes of investors dictate the trends in which prices move and they are often closely related to a variety of economic, monetary, political, and psychological forces (Pring 1991:2). New attitudes may also emerge in a financial market with sufficient stimuli, rapidly changing the way in which investors evaluate stocks and often playing a critical role in the pricing of securities. These styles often seem completely unfounded in hindsight, but can be overwhelmingly compelling to the populous at the time.

Financial markets had recently been subjected to changing attitudes towards technology sector stock, in addition to bull[11] and bear[12] markets in general. Many believed, against all previous trends, that the North American markets were in a permanent bull market, never to come down. This was, of course, before the "tech crash" of 2000. In spite of numerous similar instances, many believed that this technology trend was unprecedented and had no reason to change direction. It was said that, unlike the "dot-com" markets of a few years earlier, the rise of the technology sector was based on tangible, saleable, products that were, or would be, in high demand. True or not, investors got caught-up with the idea of making huge profits in a *sure thing* stock market, and each new buyer believed, as did the last, that *now* was the time to take advantage of the price before it went up even higher. This emotional urgency of having to act *now* and wanting not to be left behind, brings about a willingness to pay a premium to participate in what investors believe to be a sound investment. Even those who did not have faith in the technology companies themselves saw an opportunity for large profits. They, often rightly, believed that even though a stock might be highly over-valued, there are numerous other investors willing to pay an even higher price. Therefore, in a frenzied market, even over-valued

[11] A bull market represents an upwards moving market tend.
[12] A bear market represents a downwards moving market tend.

stocks presented a *sound* investment. For obvious reasons, this phenomenon is often described as the *Greater Fool* theory (Malkiel 1999:32) and it is quite common for investors to disregard foundational economic theory in favour of building *castles in the sky* out of their hopes and dreams (Malkiel 1999:85).

By nature, people are emotional and often make decisions based primarily on their feelings. For investors to buy a certain stock, or type of stock, they must find a compelling reason to do so. These reasons are usually based on a company's financial records, their product or service offerings, what other companies are doing, and what other people are doing. All of these reasons require the investor to speculate what will happen in the future.

By contrast, STFT does not rely on predicting future actions. Rather, the STFT relies on predicting the results of current actions by means of analyzing current trends or movements. There is a significant difference between predicting future actions and predicting the results of current actions. Accurately predicting future actions requires some method of achieving clairvoyance. Accurately predicting the results of current actions requires a process for the recognition of a trend before it has come to actualization. This is commonly refereed to as *pattern recognition*. Pattern recognition is used in situations in which it is necessary to determine if there is repeating thematic data. Pattern recognition is concerned with current and historic values, not necessarily the events that caused the values. There is, in fact, no need to be concerned with the current action itself, only the trend that has occurred as a result of the action. Because there is no need to attempt to input future, or even current, events, all that is needed as inputs are current and past market values for comparison. Theoretically, the greater the depth and the breadth of the historical data, the greater the efficiency and accuracy that can be expected of the forecast. This is due to the ability to develop a more complete, and therefore, accurate model. Pattern recognition, and its implementation will be discussed in detail throughout section 4.

There appears currently to be sufficient available technology to provide the STFT with a means of implementation. Moreover, the technology needed is readily available and is relatively inexpensive in comparison to the minicomputers that are often needed for other computer aided forecasting techniques. Additionally, with computer processing technology increasing at its astounding current rate, any deficiencies in currently available technology will most likely be addressed in the near future. While the availability of sufficient technology does not contribute to the accuracy of a theory, it does provide the initiative for further research into the implementation of the theory, due to the potential for immediate profitability.

Profitability is quite likely when using the STFT, as it should provide a better-than-random forecast. A completely random short-term method of investing, assuming sufficient volume and diversification[13], should produce a net return of zero averaged over time, as roughly half of the investments will go up, and roughly half of the investments will go down, in the short-term, all else being equal. Any additional accurate insight will, theoretically, move the ratio in favour of profitability[14]. The additional information makes the forecast a truly non-random forecast. There is a mathematical advantage of the forecast that makes its accuracy, and can therefore be relied upon with some degree of certainty. A forecast will only ever be as accurate as the model on which it is based, but the STFT suggests the possibility of an accurate short-term forecast, when most other theories refute it.

---

[13] Volume refers the quantity of stock within an individuals assets, or portfolio. Diversification refers to the number of different stocks held at any one point in time. With insufficient volume and/or diversification, there is a likelihood of producing a saturation of investments that do not gain in value and which diminish in value. It is for the same reason that when research data is collected, there must be a large enough data pool from which the desired data is to be extracted and the volume of data extracted must be at sufficient levels to produce meaningful results.

[14] Profitability can only be assumed when there are no additional costs involved that might diminish any gains from the trade. Two such costs are brokerage fees and commissions. However, if current trends continue, both these costs may soon be a negligible amount, and therefore less of a consideration. It should also be noted that brokerages themselves often engage in trading and are not subjected to the same fees as the common public.

*****

3.3 How does the STFT overcome the problems of traditional forecasting?

It is now appropriate to discuss how the STFT proposes to overcome the limitations of existing financial forecasting techniques, as discussed in section 3.1.

As was acknowledged at the beginning of section 3.1, it is impossible to reliably predict all the future actions of any company. Any forecasting theory attempting to reliably and correctly predict the future actions of a company is doomed to fail, as neither humans nor computers are able to consistently and accurately predict the actions of others. This is especially true when dealing with immediate and short-term events, because of the shortened allowable reaction time. A shortened allowable reaction time means that the needed research for any “action based” forecasting theory would not be able to be completed.

The STFT is *not* an “action based” forecasting theory. The STFT relies on pattern recognition based on the results of actions by others. It may be considered a “reaction based” forecasting theory. A *reaction based* forecasting system, such as the STFT, does not rely on predicting the initial actions of any company, it is only concerned with predicting the reactions, based on pattern recognition, to an initial unknown action.

An *action based* forecasting theory must rely on interpreting large volumes of abstract data. This abstract data may include such non-quantifiable market effects as emotion, weather, or seemingly unrelated market pressures. The cost associated with attempting to decipher the mammoth quantity of abstract data would far outweigh the benefits of a subsequent forecasting system. By contrast, the STFT relies only on concrete values; values such as historic stock prices. By having a knowledge-base populated with all past trades or past prices, recorded at regular short-term intervals, historic market values can be compared to current values in an attempt to find a trend. All data needed for the STFT to be implemented in a stock market is publicly and easily available from established trading markets. Furthermore, the data needed for the STFT is available instantaneously through high-speed, direct, digital electronic market connections. By contrast, an action based forecasting system must generate its own data based on user input.

With a dynamic, historic knowledge-base, a fully automated STFT-based forecasting system is arguably a possibility. A knowledge-base is a collection of all the information needed to make a forecast. It may contain raw data, relationships, and/or predetermined functional rules. A dynamic-knowledge-base differs from a static one, in that the information held within the system, including rules on which the model is based, may be altered in reaction to the current situation. The role of the knowledge-base and the dynamic-knowledge-base within a forecasting system will be explained further in subsequent sections. Because of the lack of subjectivity and human intervention needed by the STFT, all the input and output values are directly and easily quantifiable, and therefore can be manipulated mathematically. A largely automated forecasting system should be possible due to the STFT's mathematical nature. Any purely mathematical procedure may be considered a mechanical procedure and any mechanical procedure can be recorded, and therefore implemented, or programmed, through a series of precise actions, comprising an algorithm. Once an algorithm such as this is constructed, it can be repeated unsupervised.

A major concern to the financial forecasting community is the Efficient Market Hypothesis. As explained earlier, the Efficient Market Hypothesis, also known as the Random Walk Theory, hypothesizes that the speed at which the market adjusts to new information is of such a high rate that it is impossible to take systematic and sustained advantage of any market change predicted by any forecasting method. The Efficient Market Hypothesis assumes that all investors on average have exactly the same information, the same interpretation of that information, and the same emotional response to that information. Once new information is made publicly available, there is therefore a near-instantaneous market adjustment. The STFT, as discussed previously, does not rely on human interpretation of the data, nor on any emotional stimuli. If the EMH is true and markets are indeed efficient, in order for the STFT to be profitable its implementation must be very fast, as a slow implementation would result in market pressures from human forecasters overriding the initial trend, rendering it useless. This work does not intend to challenge the Efficient Market Hypothesis, only point out that it is not a factor with this method of forecasting.

The data generated in a traditional forecasting method is open to interpretation according to the views and style of the operator or programmer. When using the STFT there is limited possibility of subjectivity as compared to other forecasting methods. The process, being a simple one, is very finite, where current trends are mathematically compared to historic data. When properly implemented, there is little or no opportunity for human intervention, and therefore limited likelihood for human error. This means that there is virtually no need for highly trained and skilled experts to implement and monitor the forecasting, resulting in considerable cost saving opportunities. The profitability of a STFT based forecasting system is further increased due to its propensity for automation. The automation of a forecasting system allows for virtually unsupervised, twenty-four-hour-a-day operation.

These aspects of automation, non-supervision and continuous operation, provide two distinct and important opportunities for profitability. With no need for constant supervision, there is no cost associated with employee compensation, as an employee in a supervisory role is not necessary. An automated forecasting system also allows for continuous operation. Continuous operation provides extended hours of operation, which, accordingly, increases the opportunity for greater profitability compared to the operation hours possible with human implementation. In addition there is little need to provide costly training in order to implement and monitor the forecasting system.

With a traditional system of forecasting there is a need for very regular supervision, even after implementation. With an automated system, the need for regular supervision is greatly reduced. Through the use of a learning system, such as a neural network, discussed in further detail in subsequent sections, and a dynamic-knowledge-base, it is possible to implement an almost completely self-maintaining forecasting system. The ability of the forecasting system to learn, and update itself accordingly, allows for the most accurate and up-to-date model possible. With the passing of time, the result of a self-maintaining, learning, forecasting system will be increasing system accuracy as it builds a greater wealth of knowledge from which to draw. Compared to a system that uses static financial models, the error correction made possible by the use of a dynamic-knowledge-base system ensures the possibility of far greater accuracy.

Whilst this thesis is primarily concerned with the theoretical aspects of the STFT, it should be mentioned that the STFT also has real-world practicality. The implementation of a computerized forecasting system, based on the STFT, would admittedly take a significant amount of mathematical processing power in order to compute forecasts based on new and historical data instantaneously, or in what is called *real-time*. However, given the constant and continuing exponential increase in affordable computing power, it is likely that the STFT could be implemented using readily available technology. This topic will be revisited in subsequent sections pertaining to the implementation of such a system.

The STFT is a theory which is arguably applicable to all, or at least most, short and immediate term forecasting scenarios. Although short-term financial forecasting, in particular, may be controversial, there is no single element of the

STFT which should be considered unusual or absurd. Recall that the Short-Term Forecasting Theory states that:

> It is theoretically possible for **immediate** and **short-term** predictions to be made using historical time-series data for the reason that if a pattern of events exist between multiple specimens, it is unnecessary to know the initial cause of the pattern in order to exploit it, using pattern recognition systems. It is possible to make immediate and short-term financial predictions based on historical data, as pattern recognition is based on the *results* of the pattern instigator, not the instigator itself. *Therefore, the instigator need not be known; only the initial results.*

The STFT can be considered universally acceptable because it does not claim that everything can be predicted, only those things that have patterns in past behaviour from which future events can be extrapolated. Implementation of the STFT in areas other than financial forecasting, such as weather forecasting, political forecasting, and demographic forecasting for example, may not always be easily obtainable due to any number of constraints, including available technology, time constraints, and financial limitations. However, this does not indicate a flawed theory, only that it may be difficult to implement in certain situations. If the subject of the forecast is quantifiable, and there are patterns to be found, the STFT should be applicable.

The STFT may be applied to many different situations where short-term forecasting is of interest. Weather is a prime example of a system that can be predicted by using historical data. Quantifying weather may prove to be difficult, but if this problem could be overcome, it would be possible to predict short-term weather events, accurately (better than randomly), by using the STFT premise. One would need to compare current weather data with historical weather data in order to find a recurrent pattern. If, for example, it is consistently found that within 12 hours of a lightning of storm, in a certain location, at a certain time of year, the temperature drops 10 degrees Celsius, a forecasting system could, with a useful degree of accuracy, predict a temperature drop within 12 hours after this particular weather condition.

In addition, the complexity of a theory plays a role in its transference ability and its general acceptability. Even if a theory or method is universally applicable, if its complexity is of such a large magnitude, it may be rendered useless because of its inaccessibility. The STFT is a very simple theory, can be broken down into two generalized statements:

1. If there are recurrent relations within multiple data series, accurate predictions can be made by comparing the current and historic data of the concerned relations.
2. Predictions from statement #1 are possible because the *results* of initial actions are all that this method of forecasting is concerned with, not the initial action itself, which does not need to be predicted.

As explained earlier, it is preferable to have a dynamic knowledge-base, over a static one, for the reason that the knowledge-base is then able to develop a more complete, and therefore, accurate model. Another reason for the use of a dynamic knowledge-base is its likelihood of consistency. The accuracy of a static model may decrease over time and may result in inconsistent rates of return. This is due to a constantly changing and evolving market place. If a model is able to change and evolve with the market, it is very likely to achieve a superior level of accuracy, thereby allowing for a higher degree of precision in experimentation. Such a model allows for an experimenter to change one aspect of the program with all else remaining unchanged, review the results, and observe whether or not there has been an increase or a decrease, in terms of forecasting accuracy.

A final argument against the practicality of short-term financial forecasting is the inability to account for and track innumerable, constantly changing relevant variables. This thesis supports the popular notion that it is impractical to attempt to account for and monitor every applicable variable in short-term forecasting. The STFT only focuses on a select number of variables. These variables are initially selected by the operator based on certain assumptions and predictions about what is being forecasted, such as what quantifiable variables might hold some relation to the variable being forecasted. In terms of a stock market, every stock could be considered as a variable. The data obtained from these stock values would be compared and searched for recurrent patterns. There is no need to account for all variables if any data in excess of the stock value data is redundant. When a recurrent pattern is found, there is no need to reaffirm this pattern with numerous external variables. For this reason, one need not take into consideration all possible variables, but only those easily obtainable. Naturally, however, the greater the amount of available data from different variables, the more complete the model, which would therefore increase the potential for accuracy. However, as the number of variables increase, the number of comparisons between variables increase exponentially, requiring significantly more computing power.

*****

3.4 What are some of the limitations of the STFT?

As with every theory, the STFT has its limitations. As is suggested by its name, the Short Term Forecasting Theory in no way claims to be able to forecast medium or long-term trends. The theory only applies to short and immediate-term forecasting. The reason for this is that it remains impossible to accurately foresee the future actions of any individual company. Making an accurate and successful long-term forecast requires a forecaster to anticipate overall market trends, business decision, political and social change, and much more, well in advance of them happening. Much of this type of forecasting is based on educated guessing, not pattern recognition. The STFT only intends to exploit current immediate trends due to recent actions. The attempt to predict long-term actions of an individual company or market sector, while often including highly complex mathematical operations, most often begins with assumptions and predictions based on the experience and opinion of the forecaster. This allows for a great deal of subjectivity in market predictions. Long-term forecasting is most definitely not a mechanical procedure suited for automation. The STFT, while relying on pattern recognition, is unable to foresee future events, only the immediate results of current events. For these reasons, the STFT is only applicable to short-term forecasting and not medium or long-term forecasting.

As stated earlier, in Section 3.2, the STFT should produce results that are better than average. However, it is highly improbable that *any* implementation of *any* forecasting system, the STFT included, will result in correct predictions 100% of the time. In the course of forecasting it is to be expected that there will be a number of *false-positives* in which the forecasting system will detect a pattern that either does not exist or is altered by some unforeseeable market force. It is also to be expected that there will be a number of patterns that will go undetected by the forecasting system. Fortunately, the proposed implementation of the STFT makes use of a learning dynamic knowledge-base[15], allowing the forecasting system to *learn* from its errors and improve over time. Forecasting success must be measured in relative terms compared to a random process. Any forecasting that yields a result greater than that of a random process should be considered successful.

One limitation of any financial forecasting theory, as the EMH suggests, is that it is impossible to maintain the accuracy of a forecast as the number of users of the identical implementation of the theory increases. As mentioned in Section 3.2, as individual investors assimilate their actions, market pressures increase to

[15] Please see Section 4.3 for further details on the proposed dynamic knowledge-base.

compensate for changing supply and demand. These changes result in prices going up with increased demand, or prices going down with increased supply, as would be expected. When one investor buys or sells stock, the pressures exerted in the market are an insignificant amount, and individually would not appreciably affect the price. However, as the number of investors acting in tandem increases, the greater the overall effect on the market. A problem thus arises when a large number of investors use the same forecasting technique and cause the market to react to the forecast itself, in the process destroying the initial forecast.

*****

3.5 What are the requirements of the STFT?

The STFT requires very little data, compared to traditional methods of investing – however, the data that is required is required immediately. As could be expected from a short-term forecasting system, the speed of data delivery is crucial, as is the manipulation of that data. If either the data is delivered too slowly, or the pattern recognition is performed too slowly, the system, while still being accurate, would perform poorly because the information that has been computed might no longer be usefully reflective of the current market.

Adaptability is also critical to the proper implementation of the STFT. *Historic* data is constantly being modified. Current data is only current if nothing changes over the horizon in which it is used. When things change, as they often would in seconds or minutes, current data becomes historic data, and the model upon which the forecasting system is relying must reflect the new historic data in order to be current. This can only be achieved by a system that learns with the passing of time, and has the ability to constantly update itself.

The criteria of the STFT is fairly limited, however one last foreseeable issue could arise. If absolutely everyone involved in buying and selling equities were to use an identical forecasting system, the system would inevitably fail. This is because there would be no human influence left in the pricing of stocks, and hence, nothing upon which to base a forecast. Presumably, there would be a period of *confusion*, where few or no discernable patters would exist as all stock pricing assimilated. Eventually, all stocks would reach a system equilibrium and, all else being equal, prices would not change from their equilibrium price. This would be disastrous to every involved trading exchange, as there would be relatively few reasons to trade equities that do not change in value. Fortunately, the likelihood of this scenario ever coming to fruition is infinitesimally small, due to investors drive to always increase their profits. The need to always produce a *better* system would necessarily result in the development of *different* competing systems, and the use of any different system would most likely result in price fluctuations, and therefore a stochastic market.

*****

4. STFT Implementation

Both fuzzy-logic and neural networks are key in the implementation of a short-term financial forecasting system using the STFT concept. Fuzzy-logic makes it possible to negate much of the interference in the progression of a stock trend, which can be seen on any stock chart as insignificant upward and downward movements. Fuzzy-logic can aid in the recognition of an overall upward, downward, or other specific trend. Neural networks can be used as the ‘brain’ of the forecasting system, containing all basic predefined rules about the goals of securities trading – such as, that is it better buy a stock at a low cost and sell it at a higher one – historic knowledge of securities, and all new rules defined by the system’s pattern recognition procedure. These new rules are the elements that make the forecasting system operational. They are the rules that the decision-making unit draws upon to determine a course of action when posed with the recognition of a pattern. The combination of these two computer technologies, fuzzy-logic and neural networks, appear to offer significant advantages for financial forecasting. Jointly, they allow for a forecasting system that is able to manage nonlinear stock data (the result of fuzzy-logic) and learn from current and historic data (the result of neural networks). Brought together, fuzzy-logic and neural networks are able to offer benefits to a system not available if used independently.

Prior to this point, the theory of short-term forecasting has been discussed, but not the implementation. In the sections below, the implementation of the STFT, pertaining to financial forecasting, will be discussed. Section 4 is divided into three main sections: first, an explanation of fuzzy-logic; second, an explanation of neural networks; and third, the implementation of the STFT through the use of a combination of fuzzy-logic and neural networks.

*****

4.1 The Role of Fuzzy Logic

Fuzzy-logic is an expansion of traditional logic. In traditional logic, everything is represented as either true or false. Something *is*, or it *is not*. This is how modern computers operate. A computer's central processing unit is comprised of numerous microscopic electrical switches that can be either on or off. Depending on the combination of on and off switches, there are different meanings. Traditional logic, however, is unable to deal with many situations that humans are able to manipulate and respond to with ease. For example, language is fuzzy, as word meanings are often not finite. The word 'chair' instantaneously triggers an image in people's heads. Each person's image is similar, but no two images are necessarily the same. One person may envision an armchair, while another may picture a rocking chair. The two chairs may be different, but are both recognized as chairs. If a group of people were shown a chair, there is a strong likelihood that they would all recognize it as a chair, even though their mental images of a chair might differ from the one that is presented to them. People know that there are infinite variations on what a chair is and are able to recognize a chair, even if they have never seen the particular example in question before. With traditional logic, a chair would have to be an exact replication of an individual's mental image to register as a chair. It is either a chair (the mental image), or it is not.

Figure 4-1-1 provides a representation of the development of fuzzy-logic. Logic is the mathematical manipulation of set theory, usually developed in tandem, and hence the *correspondence* in the diagram. From classical sets comes classical logic, later followed by fuzzy sets, including the development of fuzzy numbers, which lead to infinite-valued logic, ultimately leading to the development of fuzzy logic.

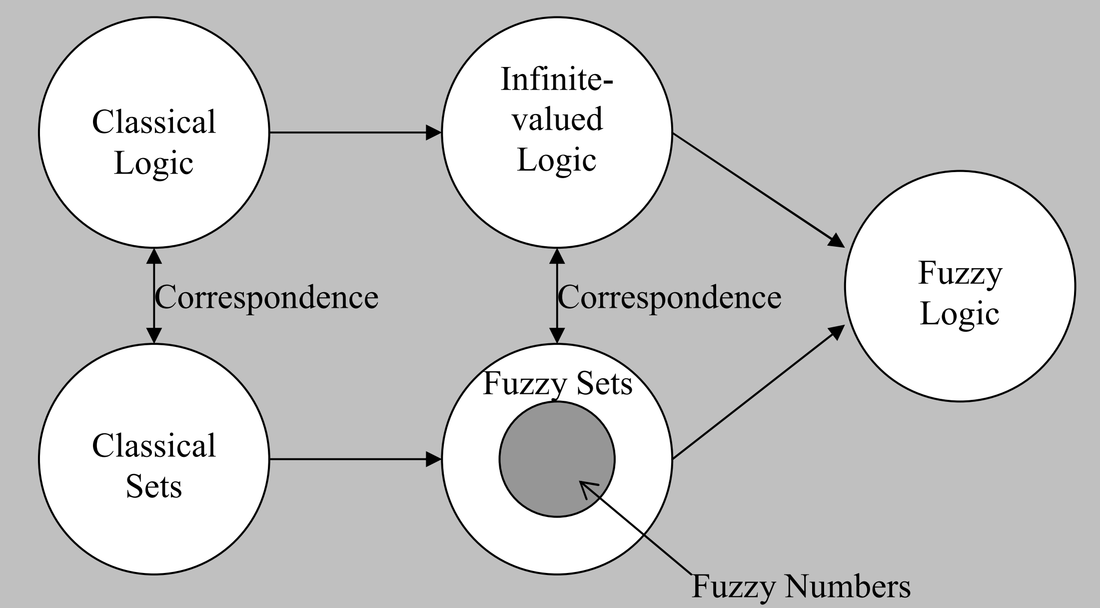

Figure 4-1-1 Development of Fuzzy Logic – adapted from Bojadzieve and Bojadzieve 1997:43

Problems, such as determining what is considered a chair, lead to what is know as set theory. When an individual reads the word 'chair', he or she understands it to be in a class of 'chair', not a particular chair. When a mathematician or logician depicts a chair as a class of furniture – or any element belonging to a class – they do so with formal models called sets (McNeill and Freiberger 1993:23). George Cantor (1845-1918) was the first to define sets. Cantor defined sets as collections of definite and distinguishable objects, real or intellectual, and that all sets are part of a universe (McNeill *et al* 1993:24). The following figure, **Figure 4-1-2**, is a pictorial representation of Cantor's various sets.

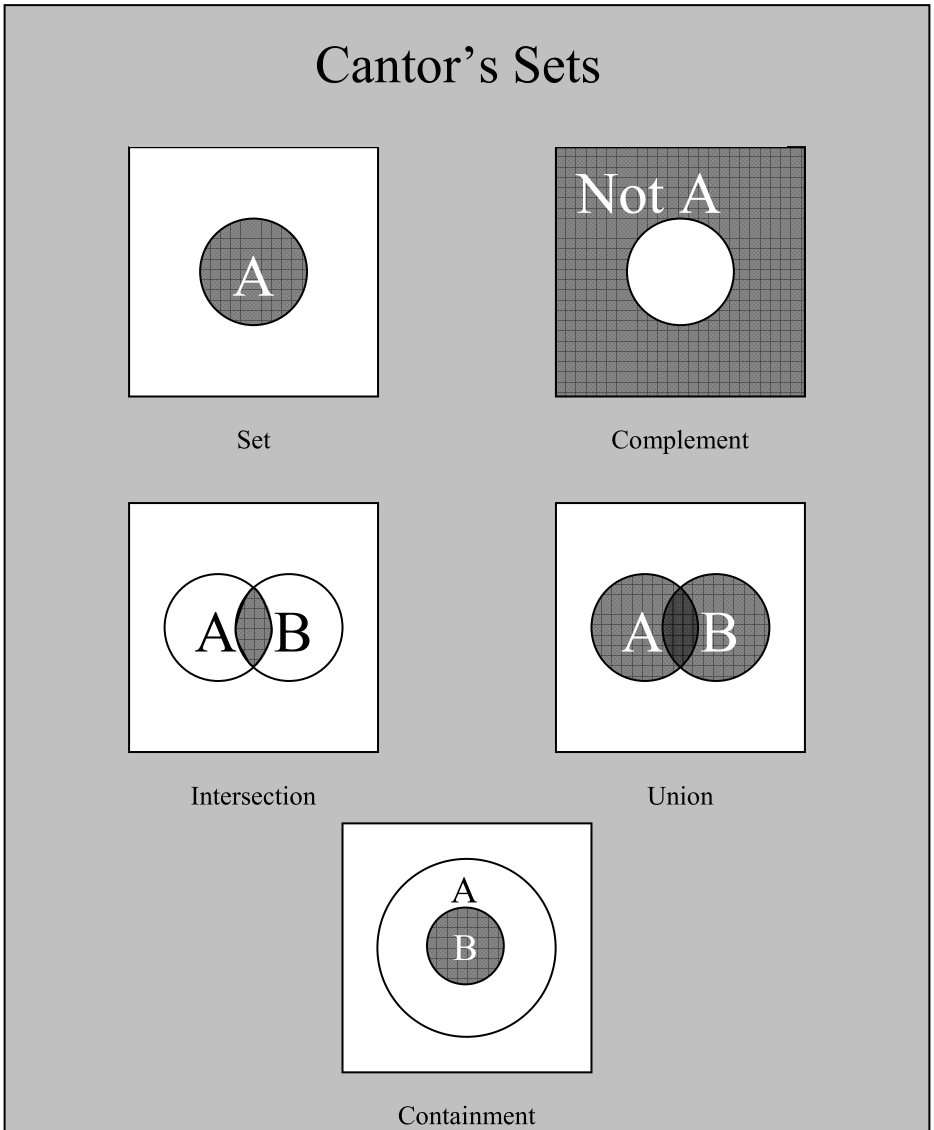

Figure 4-1-2 Cantor’s Sets–adapted from McNeill *et al* 1993:25

- The first example “Set”, is the circle A within a universe of the surrounding square. Chair is a set within a possible universe of Furniture. A “Complement” is represented by “Not A”, meaning everything within a universe that is not in set A. The complement of Chair in the first example would be the universe “Furniture, excluding any chairs”.
- An “Intersection” is the gray area that the two sets, A and B, both occupy at the same time. This is comparable to an ‘AND’ operation,

where both A and B must be true to be included in the set. Some chairs have armrests, and others have seat cushions. For example, only chairs with both armrests and seat cushions are included in an intersection in a possible universe of Furniture.

- A "Union" is the gray area contained in both sets A and B. This is as an 'OR' operation, where A or B must be true to be included. In a universe of Furniture, as above, each of 'chairs with armrests' and each of 'chairs with seat cushions' are included in a union.
- A "Containment", a variation on "Intersection" is the dark area of B, within A. Containment refers to a set that is fully within another set. Such an example would be the set 'chairs that fold' within the set 'chairs', within the possible universe Furniture.

These sets are what are called *crisp* sets. They represent TRUE and FALSE. These sets are the foundation for all mathematics concerned with logic and what is commonly referred to as the *new math*. They are also the basic operators for fuzzy-logic, although denoted differently. Cantor's set theory has been widely accepted by mathematicians, but is constantly plagued by the S*orites* ("sur-REE-tees") paradox, and similar others. The Sorites paradox is as follows:

> Take a grain of sand from a heap and you still have a heap. Take another grain from it, and it remains a heap, and so on. Eventually one grain is left. Is it still a heap? Remove it and you have nothing. Is that a heap? If not, when did it cease being one? (McNeill *et al* 1993:26-27)

Cantor's sets would have one simply decide a bright-line breakpoint, where everything beyond a certain point is no longer a heap. This is certainly counterintuitive, as the difference of one grain of sand cannot determine the existence, or not, of a heap.

The Sorites paradox is caused by vagueness. It is unclear as to when a heap is no longer a heap. Jan Łukasiewicz was the first to provide a formal model of mathematical vagueness. Łukasiewicz, consequently, developed three-valued logic. The traditional logic of true (1) and false (0) is supplemented by a halfway value (½). This is depicted as a truth-table in **Figure 4-1-3**. What is represented in this table are the values 1, ½, 0, and their opposite values. Classical logic holds that something IS (represented as the number 1), or it IS NOT (represented as the number 0). Accordingly, the opposite of 1 is 0, and vice versa. Łukasiewicz simply added the value ½, leaving room for vagueness.

STATEMENT	NEGATION
1	0
½	½
0	1

Figure 4-1-3 Three valued logic truth table

Although this third value seems obvious, it creates a perceived problem: when the value is ½, STATEMENT has the same value as NEGATION, or in other words True=False. This is what was first observed when Łukasiewicz presented his three-valued-logic to his peers, that it must be incorrect, as True can never equal False. What his colleagues failed to realize is that the ½ value is a partial contradiction, not a total contradiction. Just as when "a cup is half-full and half-empty, it is not both full and empty," (McNeill *et al* 1993:31).

Łukasiewicz then expanded his three-valued-logic. He concluded that there was no reason to insert only one value between 1 and 0. He could now add as many values as he wished. In fact, the number of intermediate values between 1 and 0 could be infinite. This not only allowed for the expression of an intermediate point but also the magnitude between the two extreme values, 1 and 0. For example, a cup one quarter full is 0.25 full and 0.75 empty. If 6'0" is considered *tall*, and 4'0" is considered *short*, Jack, at 5'10", is a fuzzy value of about 0.92 *tall* (the *tall* to *short* scale being 24 inches, and Jack being about 92% of tall), while Jill, standing at 5'0", is 0.5 *tall* (50% of tall). Using crisp values, Jack would be *short* or 0 *tall*, even though he is very close to being 6' in height, as would Jill even though she is ten inches shorter than Jack. Using multi valued, or fuzzy, logic allows for a much clearer communication to be achieved. While the addition of values between 1 and 0 may seem relatively minor and inconsequential, it was met with great skepticism as, "it defies over 2,000 years of formal logic and mathematics. [However,] it is also common sense," (Kosko 1999:3).

What makes fuzzy-logic useful for financial forecasting is that it is possible to input a series of values and mathematically determine *to what degree* is this stock moving up or down. While there might be periods where the value of a stock is slightly higher or lower than a normal value, by using fuzzy logic it is possible to compute if the stock is generally going up or down. Rather than focussing on the current and previous price of a stock to determine direction of change, a time-series, of a specific length, is analyzed using what are known as *fuzzy rule patches*, or *fuzzy IF-THEN patches*. What a *fuzzy rule patch* does is convert the crisp stock data into fuzzy data. Once the crisp data is converted to fuzzy data, it can then be compared with other fuzzy data. The purpose of this is not to forecast directly from these results, but rather to find recurrent trends in irregular and non-linear data. For a graphical representation see **Figure 4-1-4**.

While there may be trends between stocks in a market, standard crisp logic often does not allow for any variance between two similar trends. Except when using sophisticated and subjective statistical models or similar methods, any data points outside of the reference curve being compared will return a negative result, as the patterns do not match exactly. While graph A and graph B in **Figure 4-1-4** are different, they clearly share a common trend. Using a non-fuzzy linear, system, it would not be possible to match these two graphs as being similar. However, graph A+B represents the ability for fuzzy-rule patches to include graph B in the pattern of graph A. Graph A and graph B would represent a match in a fuzzy pattern recognition system in spite of them not being identical. Using fuzzy-logic, it is possible to categorize patterns, which differ slightly from each other, as being similar.

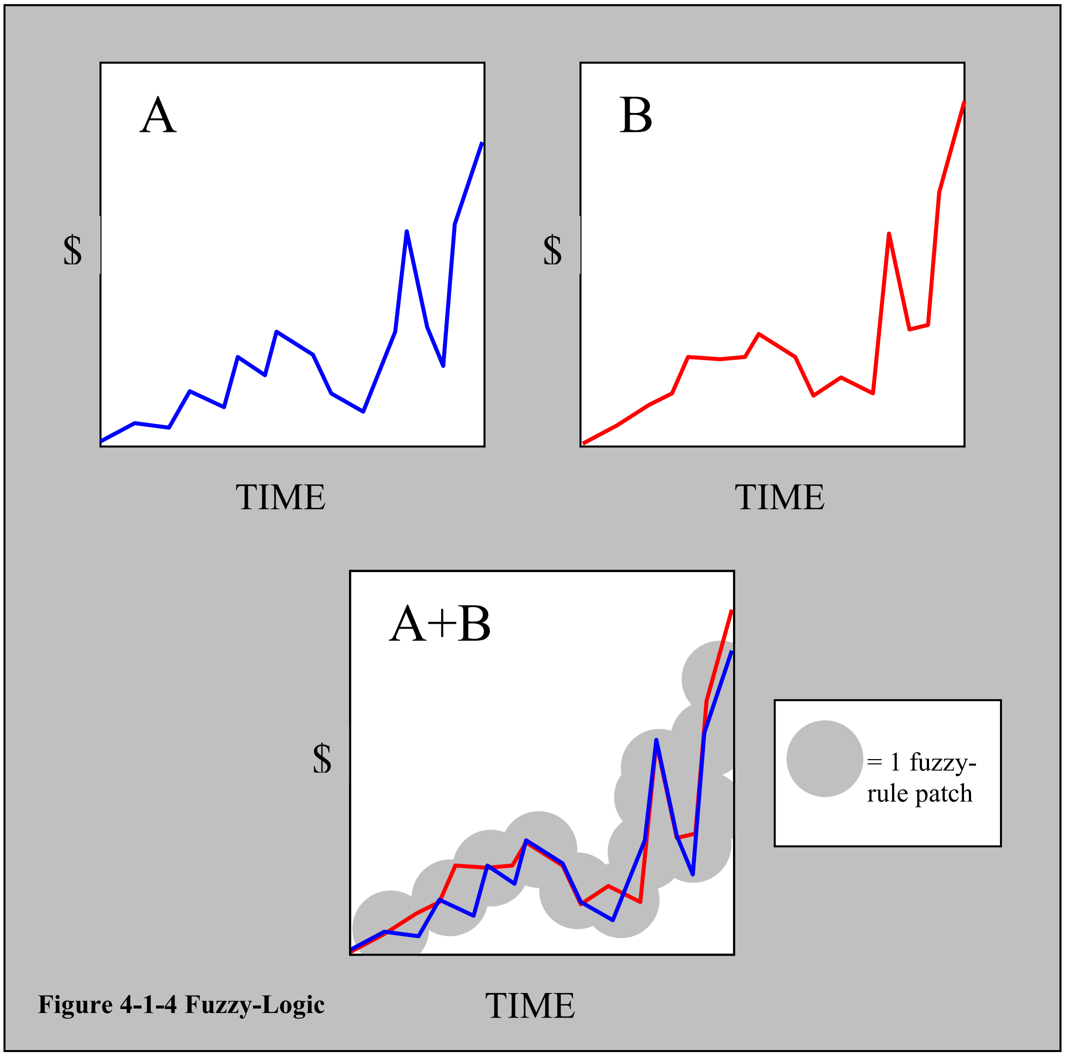

Figure 4-1-4 Fuzzy-Logic

A typical fuzzy-logic system follows the general configuration, depicted in **Figure 4-1-5**. The system is comprised of two main components: the Knowledge-Base, and the Fuzzy-Logic-Controller (FLC). The FLC is further comprised of three units: the Fuzzification Unit (Fuzz Unit), the Computation Unit, and the De-Fuzzification Unit (De-Fuzz Unit). The Fuzzification Unit receives incoming crisp data, and converts it to fuzzy data, so that it may be compared with other fuzzy data and manipulated by fuzzy operations. It does this by comparing the new data with historic data and, following rules held in the Knowledge-Base, determining to what classes the data belongs, and to what degree. The designation of the appropriate class(es), and the *degree* to which the data is a member of these classes, is then passed on to the Computation Unit. The Computation Unit performs all the necessary fuzzy computations, such as comparing the current fuzzy data to the systems overall objectives, before passing the results to the De-Fuzzification Unit.

The De-Fuzzification Unit converts the fuzzy output of the Computation Unit into crisp, usable data. While there are several methods of De-Fuzzification, the objective is to compare the data’s membership in the classes being computed, and to determining what crisp value most accurately represents the fuzzy data. The most common methods are Centre of Gravity and Centre of Maximum. The Centre of Gravity method involves determining the centre point of the function depicted by the degree of membership of the classes being compared. The Centre of Maximum method determines a crisp value by selecting the membership function with the maximum value.

The purpose of the Knowledge–Base is to provide all the necessary information, such as rules and commands, to all the various units throughout the FLC. The Knowledge-Base is where all the essential operating commands are stored. Having a Knowledge-Base as a separate unit allows for the addition, modification, and deletion of rules, without disrupting the entire system. The system may be fully functional while being modified.

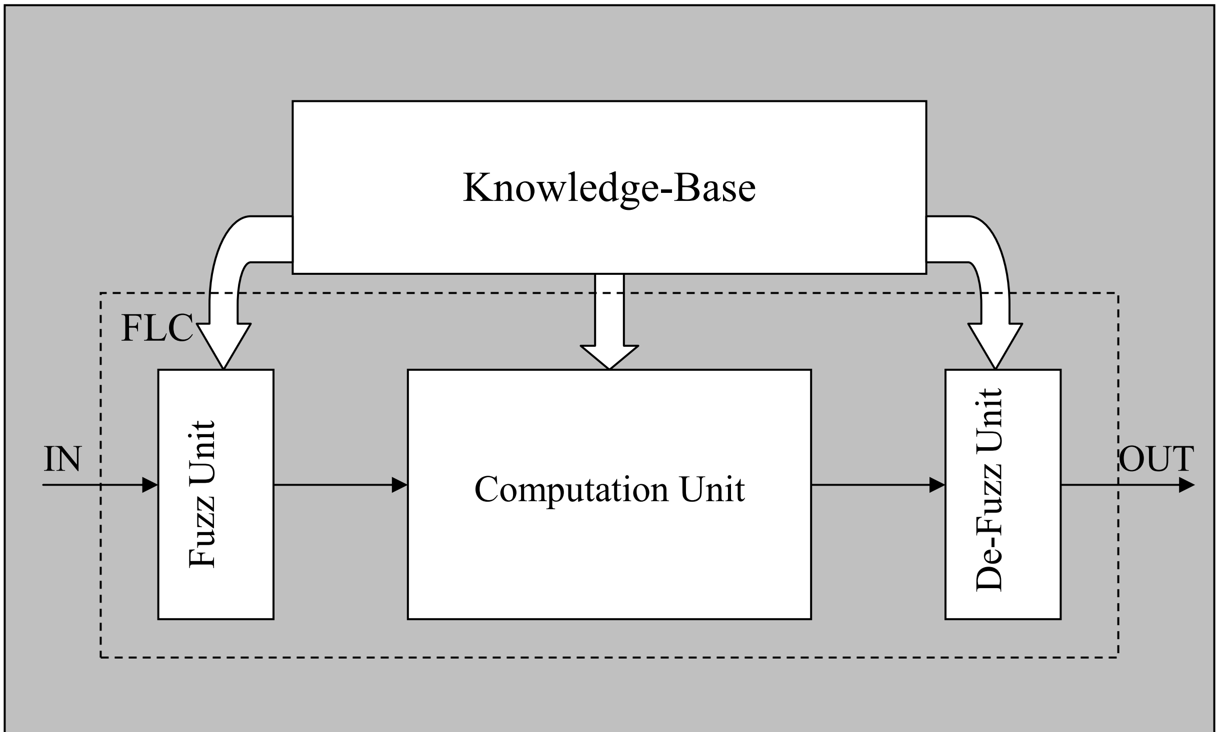

Figure 4-1-5 Fuzzy Logic System

*****

4.2 The Role of Neural Networks

A neural network, or more accurately, an artificial neural network, is a system designed to mimic the essential learning aspects of biological cerebral matter, also known as the brain. The human brain is composed of an extremely large number of neurons, each being able to complete a very simple task. When many of the neurons function together, they are able to trigger the tasks that people perform every day. An artificial neural network is a primitive model of a brain.

Both a human brain and a neural network use a method(s) of inputting data. The input devices for humans are known as senses: tactile, visual, olfactory, auditory, and gustatory. For a neural network, these inputs are digital data-streams. The inputs are processed by biological neurons in a brain, and by artificial neurons in a neural network. **Figure 4-2-1** represents a simplified neuron, called a node, as used in an artificial neural network. A basic node[16] features several input and output connections, a summation device that calculates the sum of all inputs, and an activation function to determine the node's final output (Kingdon 1997:10). In the neuron input stage, there are weighted paths that determine what is inputted to the neuron itself. Once these values are weighted and inputted, the summation device sums each inputted value and creates a representative output.

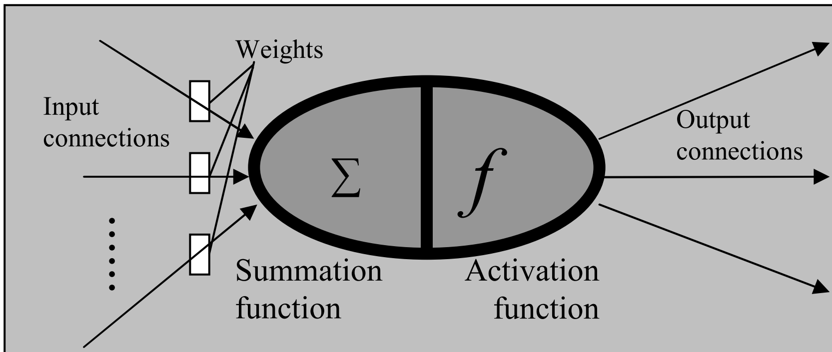

Figure 4-2-1 A conventional node– adapted from Kingdon 1997:10

Neural networks are typically organized in layers made up of a number of interconnected artificial neurons which contain a simple function that is able to pass computed data to the next neuron through weighted connections. The weights act as filters for the neuron and determine what is inputted to the neuron and subsequently what is outputted.

A neural network must be taught what constitutes a desirable outcome. For a neural network to learn, it must go through a training stage, directed by an expert. This training stage is designed to teach the neural network its ultimate

[16] For an in-depth explanation of nodes and neural networks, please see generally Kingdon (1997) and Lisboa (1992).

goals, but not how to achieve them, thereby limiting the subjectivity that is taught. A neural network is trained by providing examples of desirable outcomes, given certain inputs. As a result of learning, weighted connections are changed over time.

As the weights on the input connections increase and allow for freer data flow, the learned process becomes faster and more efficient as information travels through the path of least resistance. The neurons, biological or artificial, then pass new data to either another set of neurons or to an output stage, depending on the situation. Finally, the output stage triggers a task function of the body or computer system. **Figure 4-2-2** represents a simple neural network with four layers of neurons, of which only those called hidden layers perform data operations, while the other two, the input and output layers, simply collect and distribute pre-processed and processed data, respectively.

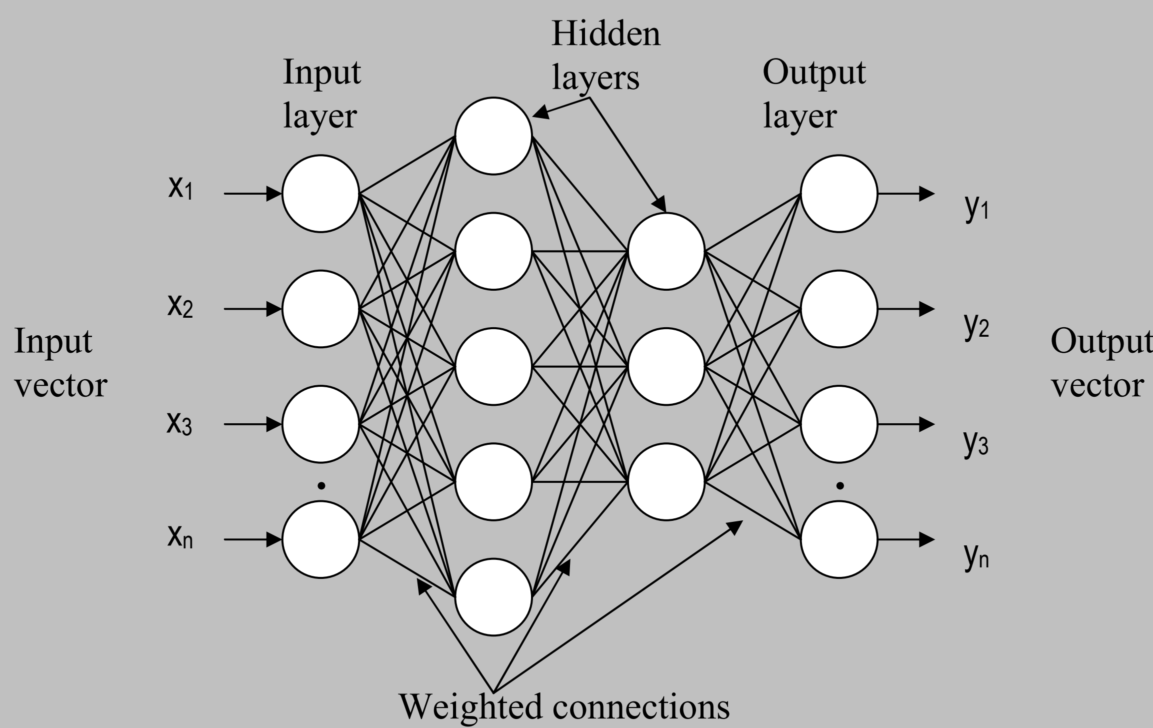

Figure 4-2-2 Four-layer Feedforward neural network – adapted from Kingdon 1997:11

Because of their learning capabilities, neural networks can be used as pattern recognition systems. In fact, pattern recognition and classification is one of the most common, and therefore proven, applications of artificial neural networks (Lisboa 1992:43). The example in **Figure 4-2-2**, called a feedforward neural network, would be suited to a pattern recognition system that does not require updating of its neurons after training. It is called a *feedforward* neural network because the network is unidirectional, where data flows straight from the input layer to the hidden layers to the output layer. There is no opportunity for data to circulate within the hidden layers, as it would in a biological neural network. There is no method for updating this particular system once the formal training stages are complete.

Other, more complicated neural networks, called *pulse-coupled neural networks*, are, however, able to maintain and update themselves. The advantage of using a more complex neural network is that, over time, not only does the network learn, but it may be designed to modify what has already been learned, in order to develop a more complete, current, and accurate model. The neurons within a pulse-coupled neural network are able to communicate with each other by way of pulses triggered by processing elements within the neurons themselves (Omidvar and Dayhoff 1998:ix). In a pulse-coupled neural network, neurons are able to communicate omnidirectionally with other neurons and, in effect, create a multi-regression neural network with the ability for self-modification. It is these pulse-coupled neural networks that seem most appropriate for financial pattern recognition. As suggested in section 3, established financial patterns may slowly change over time, eventually rendering a feedforward neural network useless for pattern recognition, because the internal model no longer reflects the current financial market. A pulse-coupled neural network would, theoretically, be able to continually modify and upgrade itself, never decreasing in accuracy, and possibly improving over time, due to a wider breadth of knowledge from which to draw.

*****

4.3 Combining Fuzzy Logic, Neural Networks and the STFT

For the STFT to be of any use, it must be implemented properly. It would appear as if the combination of fuzzy-logic and neural networks are able to meet or exceed all requirements and criteria of the STFT. That is, a system that excels at producing very fast results, accurate pattern recognition, manipulation of massive amounts of data, and the ability to modify its behaviour over time to best suit the current situation.

Figure 4-3-1 is a simplified representation of what is referred to as a *neufuz* or *neuro-fuzzy* system, a combination of a fuzzy-logic system and a neural network. A neufuz system is simply a standard fuzzy-logic system with a neural network knowledge-base. This combination allows for all the advantages of a fuzzy logic system, with the added benefits of a dynamic, learning knowledge-base. Combined, it may be possible to continually input stock data, so that the neural knowledge-base may continually learn, and continually compare a single stock value or series of stock values to the entire knowledge-base, in order to make a forecast.

A neufuz forecasting system is, theoretically, able to run virtually unsupervised, in an automated fashion. Data is inputted into both the FLC and the neural network knowledge-base. Initially, the system would not have any historical data from which to make accurate forecasts. It is thus necessary to feed the system data, and to suspend the forecasting operation. This is the formal training stage for the neural network. The result is, eventually, a knowledge-base with sufficient historical data to produce initial forecasts. The FLC is able to interpret incoming data and, by using the neural network knowledge-base, compare the current data with historic data, in an attempt to distinguish a pattern. Once a pattern is distinguished, the most suitable action is taken, determined by

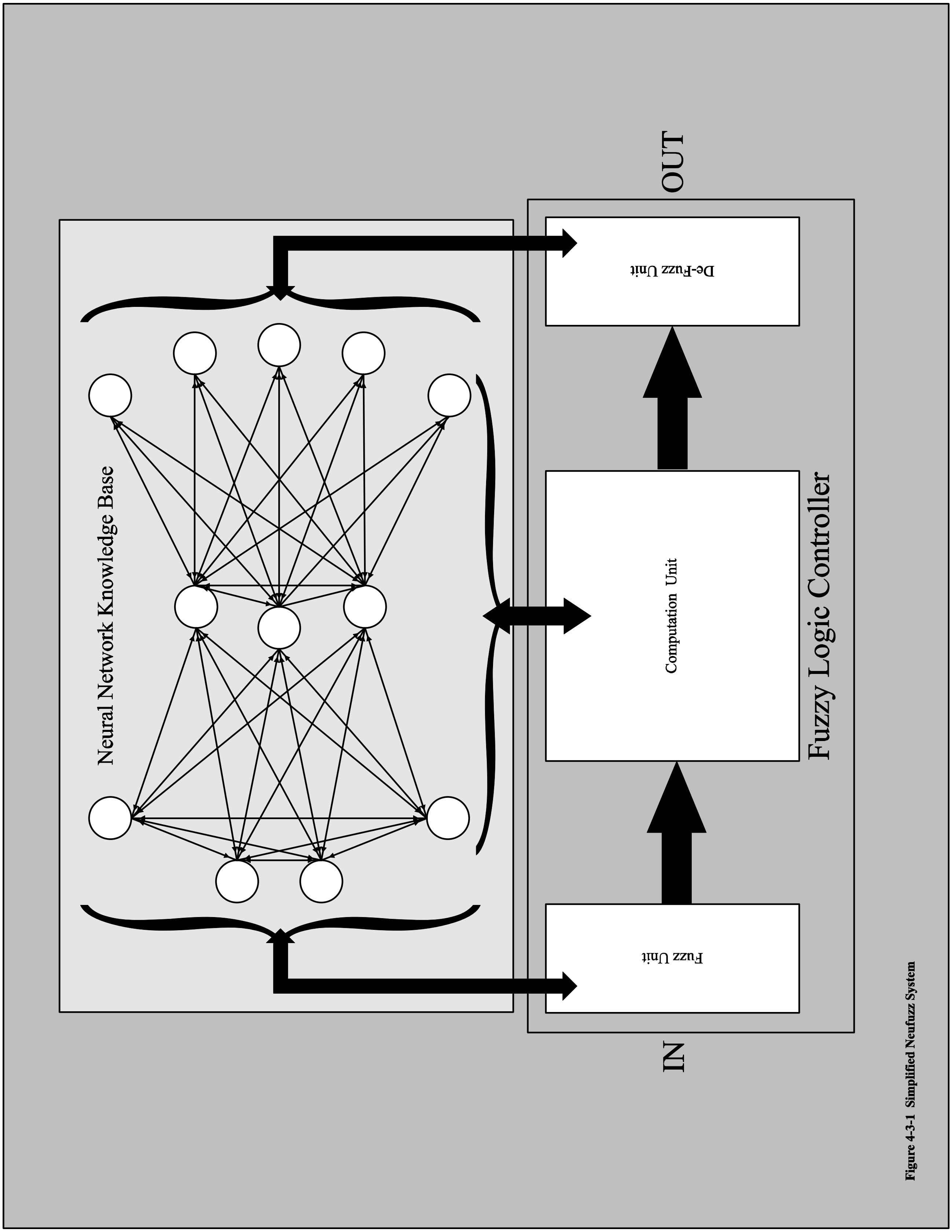

Figure 4-3-1 Simplified Neufuzz System

the neural network knowledge-base. The knowledge-base is not only able to recognize patterns, but is also able to determine the optimal timing of the reaction to a given situation. For example, it may be discovered that in order to profit from stock A going up, one must buy stock B ten minutes after the initial increase of stock A, and hold stock B for three hours, before selling. This information will be held within the neural network of patterns.

The neufuz system allows for a learning, adaptive fuzzy logic system. Fuzzy-logic, while excellent at determining the direction of a stock within a certain time, is not capable of building a knowledge-base of patterns simply by experience. For a strictly fuzzy-logic system to make accurate financial forecasts, every relevant pattern would have to be determined and programmed into the knowledge-base manually. The addition of a neural network within the fuzzy-logic knowledge-base allows for the system to learn and become adaptive. The neural network alone, while not excelling at financial forecasting, does excel at pattern recognition by learning. Together, with both the FLC and the neural network knowledge-base, it should be possible to take advantage of both systems’ strengths. The result should be a system that is able to determine the direction of a stock, to determine if a pattern is emerging and to provide forecasts of immediate short term behaviour.

*****

5. Summary

From the beginning of this paper, to the time of its completion, we have seen strong evidence of stock-market trends; trends that simply go to strengthen the validity of the STFT. We have, in all respects, seen history repeat itself. As with the Tulip-Bulb craze of Holland in 1634 to 1637 and the South Sea Bubble of 1711 in the United Kingdom, we have seen our own North American Dot-Com bubble burst with spectacular force from dizzying heights, losing as much as 78% of its value (The Nasdaq Composite). Many believed, in spite of all historic indicators, that the so-called Dot-Com market sector was to be in a perpetual upward trend. But as with all castles built in the sky, they can only float for so long, before they come crashing down. This is what happened, what *has* happened, and what *will* happen time and time again.

Markets have foreseeable trends. Markets also have foreseeable rules, such as the majority of purchases or sales of stocks are driven by fear or greed (Turner, 2000:9). As long as there is a human or natural element involved with setting the price of a stock, there will be trends to be found. People and nature are creatures of habit, and habits are, by their vary nature, foreseeable.

The promise of the STFT is not to the ability to foresee the future, nor is it some sort of ‘crystal ball’ that will tell of things to come. It is simply a theory for extrapolating data from that which already exists. The extrapolation mechanism we have been concerned with in the previous chapters is the combination of fuzzy-logic and neural networks. It is, in no way, however, to be inferred that the use of fuzzy-logic and neural networks is the only method of carrying out the STFT. Many areas of mathematics deal with the problems of pattern recognition and the process of optimization. These methods are well beyond the scope of this paper, but it should be expected that there would be more than one process that could satisfy the requirements of the STFT. The STFT says nothing of the methodology of prediction, only that a valid prediction is possible.

This paper stands to show, not that one must use fuzzy-logic and neural networks in order to properly forecast financial markets, but that the use of fuzzy-logic and neural networks is one of many possible solutions to the implementation of the STFT.

The STFT says nothing of implementation, nor does it say anything about application. The validity of a theory is strengthened by its ability to stand on its own, unfettered by a particular application. Be it financial forecasting, weather forecasting, migratory forecasting, or otherwise, the STFT claims that there are patterns in historical data that can, with an impressive degree of accuracy predict short-term future events.

*****

6. Bibliography

Abe, Shigeo. *Neural Networks and Fuzzy Systems : Theory and Applications*. Boston: Kluwer Academic, 1997.

Accounting Standards Executive Committee. *Statement of Position 74-5: Presentation and Disclosure of Financial Forecasts*. New York: American Institute of Certified Public Accounts, 1975.

American Society of Mechanical Engineers. Winter Meeting. (1992 : Anaheim Calif.), Yung C. Shin, Ahmed H. Abodelmonem, Soundar T. Kumara, and American Society of Mechanical Engineers. Production Engineering Division. *Neural Networks in Manufacturing and Robotics*. New York: American Society of Mechanical Engineers, 1992.

Auburn University. Space Power Institute., United States. National Aeronautics and Space Administration., and Society for Computer Simulation. "Proceedings of The ... Workshop on Neural Networks, Academic/Industrial/Nasa/Defense : Wnn-Aind." v. San Diego, Calif.: Society for Computer Simulation International, 1990.

Bandemer, Hans, and Siegfried Gottwald. *Fuzzy Sets, Fuzzy Logic, Fuzzy Methods with Applications*. Chichester ; New York: J. Wiley, 1995.

Bazerman, Max H. *Judgment in Managerial Decision Making*. 6th ed. Hoboken, NJ: J. Wiley, 2006.

Bengio, Yoshua. *Neural Networks for Speech and Sequence Recognition*. London ; Toronto: International Thomson Computer Press, 1996.

Biesanz, Mavis, and John Berry Biesanz. *Introduction to Sociology*. 2D Ed ed. Englewood Cliffs, N.J., Prentice-Hall (C1973), 1973.

Bishop, Christopher M. *Neural Networks for Pattern Recognition*. Oxford ; New York: Clarendon Press ; Oxford University Press, 1995.

Bojadziev, George, and Maria Bojadziev. *Fuzzy Logic for Business, Finance, and Management*, *Advances in Fuzzy Systems ; Vol. 12*. Singapore ; River Edge, NJ: World Scientific, 1997.

Chua, Jess H., and Richard S. Woodward. *Gains from Stock-Market Timing*, *Monograph Series in Finance and Economics, Monograph 1986-2*. New York, N.Y.: Salomon Brothers Center for the Study of Financial Institutions Graduate School of Business Administration New York University c1986, 1986.

Colby, Robert W. *The Encyclopedia of Technical Market Indicators, Second Edition*. New York, NY: McGraw-Hill, 2003.

DeMark, Thomas R. *The New Science of Technical Analysis*, *Wiley Finance Editions.* New York: Wiley, 1994.

Giroux, Gary A., and Peter S. Rose. *Financial Forecasting in Banking : Methods and Applications*, *Research for Business Decisions ; No. 37*. Ann Arbor, MI: UMI Research Press, 1981.

Granger, C. W. J., and Oskar Morgenstern. *Predictability of Stock Market Prices*. Lexington, Mass: Heath Lexington Books, 1970.

Greenberg, Jerald. *Behaviour in Organizations : Understanding and Managing the Human Side of Work*. 2nd Canadian ed. Toronto, Ont.: Pearson Education Canada, 2000.

Haykin, Simon S. *Neural Networks : A Comprehensive Foundation*. 2nd ed. Upper Saddle River, N.J.: Prentice Hall, 1999.

Kelejian, Harry H., and Wallace E. Oates. *Introduction to Econometrics : Principles and Applications*. 3rd ed. New York: Harper & Row, 1989.

Kingdon, J. *Intelligent Systems and Financial Forecasting*, *Perspectives in Neural Computing*. New York Springer, 1997.

Kosko, Bart. *The Fuzzy Future : From Society and Science to Heaven in a Chip*. 1st ed. New York: Harmony Books, 1999.

———. *Neural Networks and Fuzzy Systems : A Dynamical Systems Approach to Machine Intelligence*. Englewood Ciffs, NJ: Prentice Hall, 1992.

Lisboa, P. G. J. *Neural Networks : Current Applications*. London ; New York: Chapman & Hall, 1992.

Malkiel, Burton Gordon, Burton Gordon Malkiel, and netLibrary Inc. *A Random Walk Down Wall Street : Including a Life-Cycle Guide to Personal Investing*. New York: Norton, 1999.

McNeill, Dan, and Paul Freiberger. *Fuzzy Logic*. New York: Simon & Schuster, 1993.

McNeill, F. Martin, and Ellen Thro. *Fuzzy Logic : A Practical Approach*. Boston: AP Professional, 1994.

Nguyen, Hung T. *A First Course in Fuzzy and Neural Control*. Boca Raton, Fla. ; London: Chapman & Hall/CRC, 2003.

Nguyen, H. T., and E. Walker. *A First Course in Fuzzy Logic*. Boca Raton: CRC Press, 1997.

Nie, Junhong, and D. A. Linkens. *Fuzzy-Neural Control : Principles, Algorithms, and Applications*. New York ; Toronto: Prentice Hall, 1995.

Nigrin, Albert. *Neural Networks for Pattern Recognition*. Cambridge, Mass.: MIT Press, 1993.

O'Donovan, Thomas M. *Short Term Forecasting : An Introduction to the Box-Jenkins Approach*. Chichester West Sussex ; Toronto: Wiley, 1983.

Omidvar, Omid, and Judith E. Dayhoff. *Neural Networks and Pattern Recognition*. San Diego, Calif. ; London: Academic Press, 1998.

Pham, D. T., and Xing Liu. *Neural Networks for Identification, Prediction, and Control*. London ; New York: Springer-Verlag, 1995.

Pring, Martin J. *Technical Analysis Explained : The Successful Investor's Guide to Spotting Investment Trends and Turning Points*. 3rd ed. New York ; Toronto: McGraw-Hill, 1991.

Reghi÷s, Mircea, and Eugene Roventa. *Classical and Fuzzy Concepts in Mathematical Logic and Applications*. Boca Raton, Fla. ; London: CRC Press, 1998.

Ripley, Brian D. *Pattern Recognition and Neural Networks*. Cambridge ; New York: Cambridge University Press, 1996.

Ruan, Da. *Intelligent Hybrid Systems : Fuzzy Logic, Neural Networks, and Genetic Algorithms*. Boston: Kluwer Academic Publishers, 1997.

Tanaka, Kazuo. *An Introduction to Fuzzy Logic for Practical Applications*. New York: Springer, 1997.

Tanaka, Kazuo, and Hua O. Wang. *Fuzzy Control Systems Design and Analysis : A Linear Matrix Inequality Approach*. New York ; Chichester England: Wiley, 2001.

Treleaven, P. C., and IBC Technical Services Ltd. *The Third European Seminar on Neural Computing : The Marketplace : 15th & 16th February 1990, the London Marriott Hotel*. London, England: IBC Technical Services, 1990.

Turner, Toni. *A beginner’s Guide to Day Trading Online*. Avon, MA: Adams Media, 2000.

Turunen, Esko. *Mathematics Behind Fuzzy Logic*, *Advances in Soft Computing*. Heidelberg ; New York: Physica-Verlag, 1999.

Von Altrock, Constantin. *Fuzzy Logic and Neurofuzzy Applications Explained*. Englewood Cliffs, N.J.: Prentice Hall PTR, 1995.